WO 2005/096804

SEQUENCE LISTING

```
<110> Lanahan, Mike
<120> Self-processing Plants and Plant Parts
<130> 109846.317
<140> US 60/315,281
<141> 2001-08-27
<160> 112
<170> FastSEQ for Windows Version 4.0
<210> 1
<211> 436
<212> FRT
<213> Artificial Sequence
<223> synthetic
<400> 1
Met Ala Lys Tyr Leu Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala
                                  10
Phe Tyr Trp Asp Val Pro Ser Gly Gly Ile Trp Trp Asp Thr Ile Arg
                               25
                                                 30
           20
Gln Lys Ile Pro Glu Trp Tyr Asp Ala Gly Ile Ser Ala Ile Trp Ile
                          40
                                              45
Pro Pro Ala Ser Lys Gly Met Ser Gly Gly Tyr Ser Met Gly Tyr Asp
                       55
                                          60
  50
Pro Tyr Asp Tyr Phe Asp Leu Gly Glu Tyr Tyr Gln Lys Gly Thr Val
                   70
                                     75
Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Ile Asn Met Ile Asn Thr
                                   90
              8.5
Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His
                                                 110
                            105
           100
Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Gly Asp Tyr Thr
                                              125
                          120
Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr
                                         140
                      135
   130
Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe
                                  155
                150
Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp
                                  170
               165
Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly
                                                  190
                              185
          180
lle Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val
                           200
                                              205
      195
Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr
                    215
Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Ser Ser Gly
                                    235
                                                          240
                230
225
Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Ala Ala Phe
                                   250
               245
Asp Asn Lys Asn Ile Pro Ala Leu Val Glu Ala Leu Lys Asn Gly Gly
```

270

265

260

```
Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn
         275
                             280
                                                 285
  His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile
     290
                         295
                                              300
 Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu
                     310
                                         315
                                                              320
 Trp Leu Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn
                                     330
                                                          335
 Leu Ala Gly Gly Ser Thr Ser Ile Val Tyr Tyr Asp Ser Asp Glu Met
                               345
                                                      350
 lle Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr
         355
                             360
 lle Asn Leu Gly Ser Ser Lys Val Gly Arg Trp Val Tyr Val Pro Lys
                         375
                                             380
 Phe Ala Gly Ala Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp
 385
                     390
                                         395
 Val Asp Lys Tyr Val Tyr Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro
                405
                                    410
                                                        415
 Ala Tyr Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr
             420
 Cys Gly Val Gly
         435
 <210> 2
 <211> 1308
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> synthetic
<400> 2
atggccaagt acctggagct ggaggagggc ggcgtgatca tgcaggcgtt ctactgggac 60
gtcccqaqcq gaggcatctq qtqqqacacc atccqccaqa agatccccqa gtggtacqac 120
geoggeatet eegegatetg gatacegeea getteeaagg geatgteegg gggetacteg 180
atgggctacg acceptacga ctacttegae eteggegagt actaccagaa gggcacggtg 240
gagacgcgct tcgggtccaa gcaggagctc atcaacatga tcaacacggc gcacgcctac 300
ggcatcaagg teategegga categegate aaceaeaggg eeggeggega eetggagtgg 360
aaccegticg teggegacta cacetggacg gaetteteca aggicgeete eggeaagtae 420
accgccaact acctegactt ecaceccaac gagetgeacg egggegacte eggcaegtte 480
ggcggctacc cggacatetg ccacgacaag teetgggace agtactgget etgggeeteg 540
caggagteet acgeggeeta cetgegetee ateggeateg acgegtggeg ettegactae 600
gtcaagggct acggggcctg ggtggtcaag gactggctca actggtgggg cggctgggcg 660
gtgggcgagt actgggacac caacgtcgac gcgctgctca actgggccta ctcctccggc 720
gecaaggigt tegaetteee cetglactae aagatggaeg eggeettega caacaagaac 780
ateceggege tegtegagge cetgaagaac ggeggeacgg tggteteceg egaccegtte 840
aaggeegtga cettegtege caaccacgae acggacatea tetggaacaa gtacceggeg 900
tacgcettea tecteaceta egagogecag eccaegatet tetacegega etacgaggag 960
tggctgaaca aggacaaget caagaacetg atetggatte acgacaacet egegggegge 1020
tocactagta togtgtacta egacteegac gagatgatet tegteegeaa eggetaegge 1080
tecaageeeg geetgateae gtacateaae etgggeteet eeaaggtggg eegetgggtg 1140
tacgteecga agttegeegg egegtgeate cacgagtaca eeggeaacet eggeggetgg 1200
gtggacaagt acgtgtactc ctccggctgg gtctacctgg aggccccggc ctacgacccc 1260
gccaacggcc agracggcta ctccgtgtgg tcctactgcg gcgtcggc
<210> 3
```

<211> 800 <212> PRT <213> Artificial Sequence <220> <223> synthetic <400> 3 Met Gly His Trp Tyr Lys His Gln Arg Ala Tyr Gln Phe Thr Gly Glu 10 Asp Asp Phe Gly Lys Val Ala Val Lys Leu Pro Met Asp Leu Thr 20 25 30 Lys Val Gly Ile Ile Val Arg Leu Asn Glu Trp Gln Ala Lys Asp Val 35 40 45 Ala Lys Asp Arg Phe Ile Glu Ile Lys Asp Gly Lys Ala Glu Val Trp 50 60 Ile Leu Gln Gly Val Glu Glu Ile Phe Tyr Glu Lys Pro Asp Thr Ser 65 70 75 80 Pro Arg Ile Phe Phe Ala Gln Ala Arg Ser Asn Lys Val Ile Glu Ala 85 90 Phe Leu Thr Asn Pro Val Asp Thr Lys Lys Glu Leu Phe Lys Val 100 105 110 Thr Val Asp Gly Lys Glu Ile Pro Val Ser Arg Val Glu Lys Ala Asp Pro Thr Asp Ile Asp Val Thr Asn Tyr Val Arg Ile Val Leu Ser Glu 130 135 140 Ser Leu Lys Glu Glu Asp Leu Arg Lys Asp Val Glu Leu Ile Ile Glu 150 155 Gly Tyr Lys Pro Ala Arg Val Ile Met Met Glu Ile Leu Asp Asp Tyr 165 170 175 Tyr Tyr Asp Gly Glu Leu Gly Ala Val Tyr Ser Pro Glu Lys Thr Ile 180 185 190 Fhe Arg Val Trp Ser Pro Val Ser Lys Trp Val Lys Val Leu Leu Phe 195 200 205Lys Asn Gly Glu Asp Thr Glu Pro Tyr Gln Val Val Asn Met Glu Tyr 210 215 220 Lys Gly Asn Gly Val Trp Glu Ala Val Val Glu Gly Asp Leu Asp Gly 225 230 235 240 Val Phe Tyr Leu Tyr Gln Leu Glu Asn Tyr Gly Lys Ile Arg Thr Thr 245 250 255Val Asp Pro Tyr Ser Lys Ala Val Tyr Ala Asn Asn Gln Glu Ser Ala 260 265 Val Val Asn Leu Ala Arg Thr Asn Pro Glu Gly Trp Glu Asn Asp Arg 275 280 285 Gly Pro Lys Ile Glu Gly Tyr Glu Asp Ala Ile Ile Tyr Glu Ile His 295 300 Ile Ala Asp Ile Thr Gly Leu Glu Asn Ser Gly Val Lys Asn Lys Gly 310 315 320 Leu Tyr Leu Gly Leu Thr Glu Glu Asn Thr Lys Gly Pro Gly Gly Val 325 330 335 Thr Thr Gly Leu Ser His Leu Val Glu Leu Gly Val Thr His Val His $340 \hspace{1.5cm} 345 \hspace{1.5cm} 350$ lle Leu Pro Phe Phe Asp Phe Tyr Thr Gly Asp Glu Leu Asp Lys Asp 355 360 365 Phe Glu Lys Tyr Tyr Asn Trp Gly Tyr Asp Pro Tyr Leu Phe Met Val 370 375 380 Pro Glu Gly Arg Tyr Ser Thr Asp Pro Lys Asn Pro His Thr Arg Ile 390 395

Page 3

```
Arg Glu Val Lys Glu Met Val Lys Ala Leu His Lys His Gly Ile Gly
                                410
              405
Val Ile Met Asp Met Val Phe Pro His Thr Tyr Gly Ile Gly Glu Leu
                                            4 3 0
        420 425
Ser Ala Phe Asp Gln Thr Val Pro Tyr Tyr Phe Tyr Arg Ile Asp Lys
435
440
445
Thr Gly Ala Tyr Leu Asn Glu Ser Gly Cys Gly Asn Val Ile Ala Ser 450 460
  450 455
Glu Arg Pro Met Met Arg Lys Phe Ile Val Asp Thr Val Thr Tyr Trp 465 470 475 480
Val Lys Glu Tyr His Ile Asp Gly Phe Arg Phe Asp Gln Met Gly Leu
485 490 495
lle Asp Lys Lys Thr Met Leu Glu Val Glu Arg Ala Leu His Lys Ile
         500 505 510
Asp Pro Thr Ile Ile Leu Tyr Gly Glu Pro Trp Gly Gly Trp Gly Ala
                         520
                                            525
      515
Pro Ile Arg Phe Gly Lys Ser Asp Val Ala Gly Thr His Val Ala Ala
             535
                               540
Phe Asn Asp Glu Phe Arg Asp Ala Ile Arg Gly Ser Val Phe Asn Pro 545 550 560
Ser Val Lys Gly Phe Val Met Gly Gly Tyr Gly Lys Glu Thr Lys Ile
565 570 575
Lys Arg Gly Val Val Gly Ser Ile Asn Tyr Asp Gly Lys Leu Ile Lys
580 580
Ser Phe Ala Leu Asp Pro Glu Glu Thr Ile Asn Tyr Ala Ala Cys His
   595 600
                                  605
Asp Asn His Thr Leu Trp Asp Lys Asn Tyr Leu Ala Ala Lys Ala Asp 610 620
Lys Lys Glu Trp Thr Glu Glu Leu Lys Asn Ala Gln Lys Leu
625 630 635
Ala Gly Ala Ile Leu Leu Thr Ser Gln Gly Val Pro Phe Leu His Gly
                               650 655
             645
Gly Gln Asp Phe Cys Arg Thr Thr Asn Phe Asn Asp Asn Ser Tyr Asn
          660
                          665 670
Ala Pro Ile Ser Ile Asn Gly Phe Asp Tyr Glu Arg Lys Leu Gln Phe 675 680 685
Ile Asp Val Phe Asn Tyr His Lys Gly Leu Ile Lys Leu Arg Lys Glu
690 700
                   695
His Pro Ala Phe Arg Leu Lys Asn Ala Glu Glu Ile Lys Lys His Leu
705 710 715
Glu Phe Leu Pro Gly Gly Arg Arg Ile Val Ala Phe Met Leu Lys Asp
725 730 735
His Ala Gly Gly Asp Pro Trp Lys Asp Ile Val Val Ile Tyr Asn Gly 740 745
As n Leu Glu Lys Thr Thr Tyr Lys Leu Pro Glu Gly Lys Trp As n Val755 \phantom{000}760 \phantom{000}765
Val Val Asn Ser Gln Lys Ala Gly Thr Glu Val Ile Glu Thr Val Glu
               775 780
Gly Thr Ile Glu Leu Asp Pro Leu Ser Ala Tyr Val Leu Tyr Arg Glu
```

<210> 4

<211> 2400

<212> DNA

<213> Artificial Sequence

<220>

*223> synthetic <400> 4 atgggccact ggtacaagca ccagegegee taccagttea eeggegagga egactteggg 60 aaggtggccg tggtgaagct cccgatggac ctcaccaagg tgggcatcat cgtgcgcctc 120 aacgagtggc aggcgaagga cgtggccaag gaccgcttca tcgagatcaa ggacggcaag 180 geogaggigt ggatacteca gggegtggag gagatettet acgagaagee ggacacetee 240 cogogoatet tettegecca ggcccgetec aacaaggtga tegaggeett ceteaceaac 300 ccggtggaca ccaagaagaa ggagctgttc aaggtgaccg tcgacggcaa ggagatcccg 360 giglecogeg tggagaagge egaccegace gacategacg tgaccaacta egtgegeate 420 gigeteteteg agiceeteaa ggaggaggae eteegeaagg aegiggaget gaicaiegag 480 ggctacaage eggeeegeet gateatgatg gagateeteg acgactacta etacgacge 540 gagetggggg eggtgtacte eccaggagaag accatettee gegtgtggte eccggtgtee 600 aagtgggtga aggtgctcct cttcaagaac ggcgaggaca ccgagccgta ccaggtggtg 660 aacatggagt acaagggcaa cggcgtgtgg gaggccgtgg tggagggcga cctcgacggc 720 gtgttctacc tctaccaget ggagaactac ggcaagatec gcaccaccgt ggacccgtac 780 tecaaggeeg tgtaegeeaa caaccaggag tetgeagtgg tgaacetege cegeaceaac 840 coggagget gggagaacga cogcggcccg aagatcgagg gctacgagga cgccatcatc 900 tacgagatec acategeega cateacegge etggagaact eeggegtgaa gaacaaggge 960 otetaceteg decteacega ggagaacace aaggeeeegg geggegtgae caceggeete 1020 teccaceteg tggagetgeg egtgaceeae gtgcacatee tecegttett egacttetae 1080 accggcgacg agctggacaa ggacttcgag aagtactaca actggggcta cgacccgtac 1140 ctetteatgg tgeeggaggg cegetactee accgaecega agaaccegea caccegaatt 1200 cgcgaggtga aggagatggt gaaggccctc cacaagcacg gcatcggcgt gatcatggac 1260 atggtgttcc cgcacaccta cggcatcggc gagetgtccg cettegacca gaccgtgccq 1320 tactactict acceptatega caagacegge gectacetea acgagteegg etgeggeaac 1380 gtgatcgcct ccgagcgccc gatgatgcgc aagttcatcg tggacaccgt gacctactgg 1440 gtgaaggagt accacatcga eggetteege ttegaceaga tgggeeteat egacaagaag 1500 accatgotag aggragagog ogcotocac aagatogaco ogaccatoat cototacggo 1560 gageegtggg geggetgggg ggeeeegate egetteggea agteegaegt ggeeggeaec 1620 cacqtqqccq ccttcaacqa cqaqttccqc qacqccatcc qcqqctccqt qttcaacccq 1680 teegtgaagg gettegtgat gggeggetae ggeaaggaga ceaagateaa gegeggegtg 1740 gtgggeteca teaactacga eggeaagete ateaagteet tegecetega eeeggaggag 1800 accatcaact acgocgootg coacgacaac cacaccotot gggacaagaa ctacotogoo 1860 gccaaggccg acaagaagaa ggagtggacc gaggaggagc tgaagaacgc ccagaagctc 1920 geoggegeea tectecteae tagteaggge gtgccgttee tecaeggegg ceaggaette 1980 tgccgcacca ccaacttcaa cgacaactcc tacaacgccc cgatctccat caacggettc 2040 gactacgage geaageteea gitteategae gigtteaact accacaaggg ceteateaag 2100 ctocqcaagg agcaccoggo ottocgooto aagaacgoog aggagatoaa gaagcacotg 2160 gagitected egggeggeg eegealegtg geetteatge teaaggacea egeeggegge 2220 gacccgtgga aggacatcgt ggtgatctac aacggcaacc tggagaagac cacctacaag 2280 ctcccggagg gcaagtggaa cgtggtggtg aactcccaga aggccggcac cgaggtgatc 2340 gagacogtgg agggcaccat cgagctggac cogctctcog cotacgtgct ctaccgcgag 2400 <210> 5 <211> 693 <212> PRT <213> Sulfolobus solfatarious <400> 5 Met Glu Thr Ile Lys Ile Tyr Glu Asn Lys Gly Val Tyr Lys Val Val I Ile Gly Glu Pro Phe Pro Pro Ile Glu Phe Pro Leu Glu Gln Lys Ile

Page 5

25

4.5

Ser Ser Asn Lys Ser Leu Ser Glu Leu Gly Leu Thr Ile Val Gln Gln 40

Gly Asn Lys Val Ile Val Glu Lys Ser Leu Asp Leu Lys Glu His Ile

Ile Gly Leu Gly Glu Lys Ala Phe Glu Leu Asp Arg Lys Arg Lys Arg 75 70 Tyr Val Met Tyr Asn Val Asp Ala Gly Ala Tyr Lys Lys Tyr Gln Asp 85 90 95 Pro Leu Tyr Val Ser Ile Pro Leu Phe Ile Ser Val Lys Asp Gly Val Ala Thr Gly Tyr Phe Phe Asn Ser Ala Ser Lys Val Ile Phe Asp Val 115 120 125 Gly Leu Glu Glu Tyr Asp Lys Val Ile Val Thr Ile Pro Glu Asp Ser 130 140 Vai Glu Phe Tyr Val Ile Glu Gly Pro Arg Ile Glu Asp Val Leu Glu 145 150 155 Lys Tyr Thr Glu Leu Thr Gly Lys Pro Phe Leu Pro Pro Met Trp Ala 165 170 175 Phe Gly Tyr Met Ile Ser Arg Tyr Ser Tyr Tyr Pro Gln Asp Lys Val 180 185 190 Val Glu Leu Val Asp Ile Met Gln Lys Glu Gly Phe Arg Val Ala Gly 195 200 205 195 200 Val Phe Leu Asp Ile His Tyr Met Asp Ser Tyr Lys Leu Phe Thr Trp 215 220 His Pro Tyr Arg Phe Pro Glu Pro Lys Lys Leu Ile Asp Glu Leu His 230 235 Lys Arg Asn Val Lys Leu Ile Thr Ile Val Asp His Gly Ile Arg Val 245 250 Asp Gln Asn Tyr Ser Pro Phe Leu Ser Gly Met Gly Lys Phe Cys Glu 265 270 Ile Glu Ser Gly Glu Leu Phe Val Gly Lys Met Trp Pro Gly Thr Thr 275 280 285 Val Tyr Pro Asp Phe Phe Arg Glu Asp Thr Arg Glu Trp Trp Ala Gly 290 295 300 295 Leu Ile Ser Glu Trp Leu Ser Gln Gly Val Asp Gly Ile Trp Leu Asp 305 310 315 Met Asn Glu Pro Thr Asp Phe Ser Arg Ala Ile Glu Ile Arg Asp Val 325 335 Leu Ser Ser Leu Pro Val Gln Phe Arg Asp Asp Arg Leu Val Thr Thr 340 345 350 Phe Pro Asp Asn Val Val His Tyr Leu Arg Gly Lys Arg Val Lys His 355 360 365 Glu Lys Val Arg Asn Ala Tyr Pro Leu Tyr Glu Ala Met Ala Thr Phe 370 380380 Lys Gly Phe Arg Thr Ser His Arg Asn Glu Ile Phe Ile Leu Ser Arg 390 395 Ala Gly Tyr Ala Gly Ile Gln Arg Tyr Ala Phe Ile Trp Thr Gly Asp 405 410 415 Asn Thr Pro Ser Trp Asp Asp Leu Lys Leu Gln Leu Gln Leu Val Leu 425 Gly Leu Ser Ile Ser Gly Val Pro Phe Val Gly Cys Asp Ile Gly Gly 440 Phe Gln Gly Arg Asn Phe Ala Glu Ile Asp Asn Ser Met Asp Leu Leu 455 Val Lys Tyr Tyr Ala Leu Ala Leu Phe Phe Pro Phe Tyr Arg Ser His 465 470 475 480 Lys Ala Thr Asp Gly Ile Asp Thr Glu Pro Val Phe Leu Pro Asp Tyr 485 490 495 Tyr Lys Glu Lys Val Lys Glu Ile Val Glu Leu Arg Tyr Lys Phe Leu 505 500 510 Pro Tyr Ile Tyr Ser Leu Ala Leu Glu Ala Ser Glu Lys Gly His Pro

Page 6

```
520
Val Ile Arg Pro Leu Phe Tyr Glu Phe Gln Asp Asp Asp Met Tyr
            535
                                         540
Arg Ile Glu Asp Glu Tyr Met Val Gly Lys Tyr Leu Leu Tyr Ala Pro
545
                 550
                                       555
The Val Ser Lys Glu Glu Ser Arg Leu Val Thr Leu Pro Arg Gly Lys
                                                      575
               565
                                 570
Tro Tyr Asn Tyr Trp Asn Gly Glu Ile Ile Asn Gly Lys Ser Val Val
                             585
                                                590
           580
Lys Ser Thr His Glu Leu Pro Ile Tyr Leu Arg Glu Gly Ser Ile Ile
                           600
                                             605
Pro Leu Glu Gly Asp Glu Leu Ile Val Tyr Gly Glu Thr Ser Phe Lys
                      615
                                           620
    610
Arg Tyr Asp Asn Ala Glu Ile Thr Ser Ser Ser Asn Glu Ile Lys Phe
                   630
                                      635
Ser Arg Glu Ile Tyr Val Ser Lys Leu Thr Ile Thr Ser Glu Lys Pro
               645
                               650
Val Ser Lys Ile Ile Val Asp Asp Ser Lys Glu Ile Gln Val Glu Lys
                                               670
                             665
Thr Met Gln Asn Thr Tyr Val Ala Lys Ile Asn Gln Lys Ile Arg Gly
                           680
                                              685
        675
Lys Ile Asn Leu Glu
<210> 6
<211> 2082
<212> DNA
<213> Sulfolobus solfataricus
atggagacca tcaagatcta cgagaacaag ggcgtgtaca aggtggtgat cggcgagccg 60
ttcccgccqa tcgagttccc gctcgagcag aagatctcct ccaacaagtc cctctccgag 120
ctgggcctca ccatcgtgca gcagggcaac aaggtgatcg tggagaagtc cctcgacctc 180
aaggagcaca teateggeet eggegagaag geettegage tggacegcaa gegcaagege 240
tacgtgatgt acaacgtgga cgccggcgcc tacaagaagt accaggaccc gctctacgtg 300
recatecege tetteatete egtgaaggae ggegtggeea eeggetaett etteaactee 360
gcctccaagg tgatcttcga cgtgggcctc gaggagtacg acaaggtgat cgtgaccatc 420
ccqgaggact ccqtqqagtt ctacqtqatc gagggcccqc gcatcqagga cqtqctcqag 480
aagtacaccg agetgaccgg caagecgtte etecegeega tgtgggeett eggetacatg 540
atotocogot actoctacta coogeaggac aaggtggtgg agotggtgga catcatgcag 600
aaggaggget teegegtgge eggegtgtte etegacatee actacatgga etectacaag 660
ctetteacet ggeacecqta cegetteeeg gageegaaga ageteatega egagetgeae 720
aagegcaacg tgaageteat caccategtg gaccaeggea teegegtgga ecagaactae 780
toccogttoc totcoggoat gggcaagtto tgcgagatog agtccggcga gctgttcgtg 840
ggcaagatgt ggccgggcac caccgtgtac ccggacttct tccgcgagga cacccgcgag 900
tggtgggeeg geoteatete egagtggete teccagggeg tggaeggeat etggetegae 960
atgaacgage egacegaett etecegegee ategagatee gegacgtget etecteete 1020
coggracage teogogacga ecgeotogig accaection eggacaacgi ggigcactae 1080
ctccgcggca agcgcgtgaa gcacgagaag gtgcgcaacg cctacccgct ctacgaggcg 1140
auggecacet teaagggett eegeacetee cacegeaacg agatetteat eeteteeege 1200
geoggetacg eeggeateca gegetacgee tteatetgga eeggegacaa cacecegtee 1260
tgggacgacc tcaageteea getecagete gtgeteggee tetecatete eggegtgeeg 1320
ttcgtgggct gcgacatcgg cggcttccag ggccgcaact tcgccgagat cgacaactcg 1380
atggacetee tegtgaagta etaegeeete geeetettet teeegtteta eegeteeeae 1440
aaggccaccy acggcatcga caccgagccg gtgttcctcc cggactacta caaggagaag 1500
gtgaaggaga tegtggaget gegetacaag tteeteeegt acatetaete cetegeeete 1560
gaggeeteeg agaagggeea eeeggtgate egeeegetet tetacgagtt eeaggacgae 1620
```

Page 7

```
gacqacatgt accgcatcga ggacqagtac atggtgggca agtacctcct ctacgccccg 1680
arcorgica aggaggagte cogectegtg acceteege geggeaagtg gtacaactae 1740
tggaacggcg agatcatcaa cggcaagtcc gtggtgaagt ccacccacga gctgccgatc 1800
tacctccgcg agggctccat catcccgctc gagggcgacg agctgatcgt gtacggcgag 1860
accreettea agegetaega caaegeegag ateaceteet eetecaaega gateaagtte 1920
tecegegaga tetaegtigte caageteace ateaceteeg agaageeggt gtecaagate 1980
ateqtqqacq actecaagga gatecaggtg gagaagacca tgcagaacac ctacgtqqcc 2040
aagatcaacc agaagatccg cggcaagatc aacctcgagt ga
                                                                   2082
<210> 7
<211> 1818
<212> DNA
<213> Artificial Sequence
<220>
<223> synthetic
<400> 7
atggcjącto tągocacąto gcagotogto gcaacgogog coggootggg cgtocoggao 60
gegreeacgt teegeegegg egeegegeag ggeetgaggg gggeeeggge gteggeggeg 120
geggacaege teageatgeg gaceagegeg egegggege ceaggeacea geaceageag 180
gagagedec gagacaggtt coagtagete gtagtgtag caagagaagg catgaacgtc 240
grettegteg gegeegagat ggegeegtgg ageaagaceg gaggeetegg egaegteete 300
ageageetae egoogaceat ageogegaac aggeaccata teatagteat eteteccaca 360
tacqaccagt acaaggacgo otgggacaco agogtogtgt cogagatoaa gatgggagac 420
eggtacqaga eggteaggtt ettecaetge tacaagegeg gagtggaceg egtgttegtt 480
gaccacccac tgttcctgga gagggtttgg ggaaagaccg aggagaagat ctacgggcct 540
gtcgctggaa cggactacag ggacaaccag ctgcggttca gcctgctatg ccaggcagca 600
cttqaaqctc caaggateet gageeteaac aacaacecat actteteegg accataeggg 660
gaggacoteg tgttegtetg caaegactgg cacaeeggee etetetegtg etaceteaag 720
agcaactace agteceacgg catetacagg gacgeaaaga cegetttetg catecacaac 780
attitionated agggeogett egectioned gactacoogg agetgaacet ecoegagaga 840
ttcaagtcgt ccttcgattt catcgacggc tacgagaagc ccgtggaagg ccggaagatc 900
aactggatga aggccgggat cctcgaggcc gacagggtcc tcaccgtcag cccctactac 960
geographic teateteegg categorage geotgegage tegacaacat categorete 1020
accogcatea ecogeategt caacogcato gacoteageo agtogoacee caocaoogae 1080
aagtacatcg ccgtgaagta cgacgtgtcg acggccgtgg aggccaaggc gctgaacaag 1140
gaggcgctgc aggcggaggt cgggctcccg gtggaccgga acatcccgct ggtggcgttc 1200
ateggeagge tggaagagea gaagggeece gaegteatgg eggeegeeat eeegeagete 1260
atggagatgg tggaggacgt gcagatcgtt ctgctgggca cgggcaagaa gaagttcgag 1320
cgcatgctca tgagcgccga ggagaagttc ccaggcaagg tgcgccgt ggtcaagttc 1380
aacqcqqqqc tqqcqcacca catcatggcc ggcgccqacg tgctcgccgt caccagccgc 1440
ttcgagccct qcggcctcat ccagctgcag gggatgcgat acggaacgcc ctgcgcctgc 1500
acquedaccq qtqqactcqt cgacaccate atcgaaggca agaccgqqtt ccacatgqqc 1560
egecteageg tegactgeaa egtegtggag eeggeggaeg teaagaaggt ggecaceace 1620
ttgcagcgcg ccatcaaggt ggtcggcacg ccggcgtacg aggagatggt gaggaactgc 1680
argatedagg arcteteerg gaagggeest gecaagaact gggagaacgt getgeteage 1740
ctoggggteg coggeggega godaggggtt gaaggegagg agategegee getegeeaag 1800
gagaacgtgg ccgcccc
                                                                  1818
<210> 8
<211> 606
<212> PRT
<213> Artificial Sequence
<220>
<223> synthetic
```

<400> 8 Met Ala Ala Leu Ala Thr Ser Gln Leu Val Ala Thr Arg Ala Gly Leu Gly Val Pro Asp Ala Ser Thr Phe Arg Arg Gly Ala Ala Gln Gly Leu Arg Gly Ala Arg Ala Ser Ala Ala Ala Asp Thr Leu Ser Met Arg Thr 4.5 Ser Ala Arg Ala Ala Pro Arg His Gln His Gln Gln Ala Arg Arg Gly Ala Arg Phe Pro Ser Leu Val Val Cys Ala Ser Ala Gly Met Asn Val Val Phe Val Gly Ala Glo Met Ala Pro Trp Ser Lys Thr Gly Gly Leu 8.5 Gly Asp Val Leu Gly Gly Leu Pro Pro Ala Met Ala Ala Asn Gly His Arg Val Met Val Val Ser Pro Arg Tyr Asp Gln Tyr Lys Asp Ala Trp Asp Thr Ser Val Val Ser Glu Ile Lys Met Gly Asp Gly Tyr Glu Thr Val Arg Phe Phe His Cys Tyr Lys Arg Gly Val Asp Arg Val Phe Val Asp His Pro Leu Phe Leu Glu Arg Val Trp Gly Lys Thr Glu Glu Lys The Tyr Gly Pro Val Ala Gly Thr Asp Tyr Arg Asp Asn Gln Leu Arg Phe Ser Leu Leu Cys Gln Ala Ala Leu Glu Ala Pro Arg Ile Leu Ser Leu Asn Asn Asn Pro Tyr Phe Ser Gly Pro Tyr Gly Glu Asp Val Val Phe Val Cys Asn Asp Trp His Thr Gly Pro Leu Ser Cys Tyr Leu Lys Ser Asn Tyr Gln Ser His Gly Ile Tyr Arg Asp Ala Lys Thr Ala Phe Cys Ile His Asn Ile Ser Tyr Gln Gly Arg Phe Ala Phe Ser Asp Tyr Pro Glu Leu Asn Leu Pro Glu Arg Phe Lys Ser Ser Phe Asp Phe Ile Asp Gly Tyr Glu Lys Pro Val Glu Gly Arg Lys Ile Asn Trp Met Lys Ala Gly Ile Leu Glu Ala Asp Arg Val Leu Thr Val Ser Pro Tyr Tyr Ala Glu Glu Leu Ile Ser Gly Ile Ala Arg Gly Cys Glu Leu Asp Asn He Met Arg Leu Thr Gly He Thr Gly He Val Asn Gly Met Asp Val Ser Glu Trp Asp Pro Ser Arg Asp Lys Tyr Ile Ala Val Lys Tyr Asp Val Ser Thr Ala Val Glu Ala Lys Ala Leu Asn Lys Glu Ala Leu Gln Ala Glu Val Gly Leu Pro Val Asp Arg Asn Ile Pro Leu Val Ala Phe 385 390 395 400 Ile Gly Arg Leu Glu Glu Gln Lys Gly Pro Asp Val Met Ala Ala Ala lle Pro Gln Leu Met Glu Met Val Glu Asp Val Gln Ile Val Leu Leu Gly Thr Gly Lys Lys Phe Glu Arg Met Leu Met Ser Ala Glu Glu Lys Phe Pro Gly Lys Val Arg Ala Val Val Lys Phe Asn Ala Ala Leu

```
450
                        455
                                            460
Ala His His Ile Met Ala Gly Ala Asp Val Leu Ala Val Thr Ser Arg
                    470
                                       475
Phe Glu Pro Cys Gly Leu Ile Gln Leu Gln Gly Met Arg Tyr Gly Thr
                485
                                    490
                                                       495
Pro Cys Ala Cys Ala Ser Thr Gly Gly Leu Val Asp Thr Ile Ile Glu
            500
                               505
                                                 510
Gly Lys Thr Gly Phe His Met Gly Arg Leu Ser Val Asp Cys Ash Val
        515
                          520
                                                525
Val Glu Pro Ala Asp Val Lys Lys Val Ala Thr Thr Leu Gln Arg Ala
                        535
                                          540
    530
!le Lys Val Val Gly Thr Pro Ala Tyr Glu Glu Met Val Arg Asn Cys
                    550
                                       555
Met Ile Gln Asp Leu Ser Trp Lys Gly Pro Ala Lys Asn Trp Glu Asn
                                   570
                565
Val Leu Leu Ser Leu Gly Val Ala Gly Gly Glu Pro Gly Val Glu Gly
                     585
            580
Glu Glu Ile Ala Pro Leu Ala Lys Glu Asn Val Ala Ala Pro
        595
                            600
<210> 9
<211> 2223
<212> DNA
<213> Artificial Sequence
<220>
<223> synthetic
<400> 9
atggccaagt acctggaget ggaggagggc ggcgtgatca tgcaggcgtt ctactgggac 60
greecgageg gaggeatetg groggaeace atecgeeaga agateceega grogtaegae 120
geoggeatet eegegatetg gatacegeea getteeaagg geatgteegg gggetaeteg 180
atgggctacg accogtacga ctacttcgac ctcggcgagt actaccagaa gggcacggtg 240
gagacgcgct tcgggtccaa gcaggagctc atcaacatga tcaacacggc gcacgcctac 300
ggcatcaagg tcatcgcgga catcgtgatc aaccacaggg ccggcggcga cctggagtgg 360
aacccgttcg teggegacta cacctggacg gactteteea aggtegeete eggeaagtae 420
accoccaact acctogactt coaccocaac gagotgoacg cgggogacto cggcacgtto 480
ggcggctacc cggacatctg ccacgacaag tcctgggacc agtactggct ctgggcctcg 540
caggagtect acgcggccta cotgcgctcc atcggcatcg acgcgtggcg ottogactac 600
grcaaggget acggggeetg ggrggteaag gactggetea actggtgggg cggctgggcg 660
gtgggcgagt actgggacac caacgtcgac gcgctgctca actgggccta ctcctccggc 720
gccaaggtgt tegaetteec cetgtaetae aagatggaeg eggeettega caacaagaac 780
atconggege togtogagge cotgaagaac ggcggcacgg tggtctcccg cgacccgttc 840
aaqoocqtqa cottoqtoqo caaccacqao acqqacatca totqqaacaa qtacccqqcq 900
tacgcettea tecteaceta egagggecag eccaegatet tetacegega etacgaggag 960
tggctgaaca aggacaaget caagaacetg atctggatte acgacaacet cgcgggcggc 1020
tocactagta togtgtacta ogactoogae gagatgatet togtoogeaa oggotacgge 1080
tecaageeg geetgateae gtacateaae etgggeteet ecaaggtggg eegetgggtg 1140
tacqtcccqa aqttcgccqq cqcqtqcatc cacqaqtaca ccqqcaacct cqqcqqctqq 1200
gtggacaagt acqtgtactc ctccggctgg gtctacctgg aggccccggc ctacgacccc 1260
gccaacggcc agtacggcta etccgtgtgg tectaetgeg gegteggeae ategattget 1320
ggcatceteg aggeegacag ggteeteace gteageeect actaegeega ggageteate 1380
tacagoatag caagaggata agastagaa aacataatga gootcaacagg catcacagga 1440
atogicaecg gcatggacgi cagcgagigg gaccccagca gggacaagia catcgccgig 1500
aagtacgacg tgtcgacggc cgtggaggcc aaggcgctga acaaggaggc gctgcaggcg 1560
gaççıcçççc teceggiçga eeggaacate eegetggigg egiteategg eaggeiggaa 1620
gaccagaagg ccccgacgt catggcggcc gccatcccgc agctcatgga gatggtggag 1680
```

```
çacgtgcaga tegttetget gggcaeggge aagaagaagt tegagegeat geteatgage 1740
gccgaggaga agttcccagg caaggtgcgc gccgtggtca agttcaacgc ggcgctggcg 1800
caccacatca tggccggcgc cgacgtgctc gccgtcacca gccgcttcga gccctgcggc 1860
ctcatccage tgcaggggat gegataegga aegeeetgeg eetgegegte caceggtgga 1920
ctcgtcgaca ccatcatcga aggcaagacc gggttccaca tgggccgcct cagcgtcgac 1980
tgcaacqtcg tqqaqccggc ggacqtcaag aaggtggcca ccaccttgca gcgcgccatc 2040
aaggtggtcg gcacgccggc gtacgaggag atggtgagga actgcatgat ccaggatete 2100
tcctggaagg gccctgccaa gaactgggag aacgtgctgc tcagcctcgg ggtcgccggc 2160
ggcgagccag gggttgaagg cgaggagatc gcgccgctcg ccaaggagaa cgtggccgcg 2220
<210> 10
<211> 741
<212> PRT
<213> Artificial Sequence
<220>
<223> synthetic
<400> 10
Met Ala Lys Tyr Leu Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala
                                   10
Phe Tyr Trp Asp Val Pro Ser Gly Gly Ile Trp Trp Asp Thr Ile Arg
                                25
           20
Gin Lys lie Pro Glu Trp Tyr Asp Ala Gly Ile Ser Ala Ile Trp Ile
                           40
                                               45
        3.5
Pro Pro Ala Ser Lys Gly Met Ser Gly Gly Tyr Ser Met Gly Tyr Asp
                       55
    50
Pro Tyr Asp Tyr Phe Asp Leu Gly Glu Tyr Tyr Gln Lys Gly Thr Val
                                       75
                   7.0
Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Ile Asn Met Ile Asn Thr
                                   90
               85
Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His
                                                   110
                               105
          100
Aro Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Gly Asp Tyr Thr
                            120
                                                125
       115
Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr
                        135
                                           140
   130
Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe
                                       155
                   150
145
Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp
                                  170
               165
Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly
                                                   190
                                185
           180
Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val
                                              205
                            200
        195
Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr
                                          220
                        215
Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Ser Ser Gly
                                     235
                    230
Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Ala Ala Phe
                                                       255
               245
                                   250
Asp Asn Lys Asn Ile Pro Ala Leu Val Glu Ala Leu Lys Asn Gly Gly
                                265
                                                    270
Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn
                          280
                                               285
       275
His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile
                        295
                                            300
   290
```

Leu	Thr	Tyr	Glu	Gly	Gln	Pro	Thr	Ile	Phe	Tyr	Arg	Asp	Tyr-	Glu	Glu
305 Trp	Leu	Asn	Lys	Asp	310 Lys	Leu	Lys	Asn	Leu	315 Ile	Trp	Ile	His	Asp	320 Asn
			-	325	-				330					335	
		-	340		Thr			345					350		
lle	Phe	Val 355	Arg	Asn	Gly	Tyr	Gly 360	Ser	Lys	Pro	Gly	Leu 365	Ile	Thr	Tyr
lle	Asn 370	Leu	Gly	Ser	Ser	Lys 375	Val	Gly	Arg	Trp	Val 380	Tyr	Val	Pro	Lys
Phe	Ala	Gly	Ala	Cys	11e 390	His	Glu	Tyr	Thr	Gly 395	Asn	Leu	Gly	Gly	Trp 400
	Asp	LУs	Tyr	Val 405	Tyr	Ser	Ser	Gly	Trp 410	Val	Tyr	Leu	Glu	Ala 415	
Ala	Tyr	Asp	Pro 420		Asn	Gly	Gln	Tyr 425		Tyr	Ser	Val	Trp 430	Ser	Tyr
Cys	Gly	Val 435		Thr	Ser	Ile	Ala 440		Ile	Leu	Glu	Ala 445		Arg	Val
Leu			Ser	Pro	Tyr	Tyr 455		Glu	Glu	Leu	Ile 460		Gly	Ile	Ala
Arg 465	450 Gly	Cys	Glu	Leu	Asp 470		Ile	Met	Arg	Leu 475		Gly	Ile	Thr	Gly 480
	Val	Asn	Gly	Met 485	Asp	Val	Ser	Glu	Trp 490		Pro	Ser	Arg	Asp 495	
Tyr	Ile	Ala	Val 500		Tyr	Asp	Val	Ser 505		Ala	Val	Glu	Ala 510		Ala
Leu	Asn	Lys 515		Ala	Leu	Gln	Ala 520		Val	Gly	Leu	Pro 525		Asp	Arg
Asn	Ile 530		Leu	Val	Ala	Phe 535		Gly	Arg	Leu	Glu 540		Gln	Lys	Gly
Pro 545		Val	Met	Ala	Ala 550		Ile	Pro	Gln	Leu 555	Met	Glu	Met	Val	Glu 560
Asp	Val	Gln	Ιle	Val 565	Leu	Leu	Gly	Thr	Gly 570	Lys	Lys	Lys	Phe	Glu 575	Arg
Met	Leu	Met	Ser 580		Glu	Glu	Lys	Phe 585	Pro	Gly	Lys	Val	Arg 590	Ala	Val
Val	Lys	Phe 595		Ala	Ala	Leu	Ala 600		His	Ile	Met	Ala 605	Gly	Ala	Asp
Val	Leu 610		Val	Thr	Ser	Arg 615		Glu	Pro	Cys	Gly 620	Leu	Ile	Gln	Leu
Gln 625		Met	Arg	Tyr	Gly 630		Pro	Cys	Ala	Cys 635	Ala	Ser	Thr	Gly	Gly 640
Leu	Val	Asp	Thr	Ile 645	Ile	Glu	Gly	Lys	Thr 650	Gly	Phe	His	Met	Gly 655	Arg
Leu	Ser	Val	A.sp 660	Cys	Asn	Val	Val	Glu 665	Pro	Ala	Asp	Val	Lys 670	Lys	Val
Ala	Thr	Thr 675	Leu	Gln	Arg	Ala	Ile 680	Lys	Val	Val	Gly	Thr 685	Pro	Ala	Tyr
Glu	Glu 690		Val	Arg	Asn	Cys 695		Ile	Gln	Asp	Leu 700	Ser	Trp	Lys	Gly
705	Ala	-			Glu 710	Asn				715					720
Gĺy	Glu	Pro	Gly	Val 725	Glu	Gly	Glu	Glu	Ile 730	Ala	Pro	Leu	Ala	Lys 735	Glu
Asn	Val	Ala	Ala 740												

```
<210> 11
<211> 1515
<212> DNA
<213> Zea mays
<400> 11
ggagagetat gagaegtatg teeteaaage caetttgeat tgtgtgaaac caatategat 60
ctttottact tcatcatqca tqaacatttg tqqaacatac taqcttacaa qcattagtga 120
cageteaçãa aaaaçttate tatgaaaggt tteatgtgta eegtgggaaa tgagaaatgt 180
tgccaactca aacaccttca atatgttgtt tgcaggcaaa ctcttctgga agaaaggtgt 240
ctaaaactat gaacgggtta cagaaaggta taaaccacgg ctgtgcattt tggaagtatc 300
atctatagat gtctgttgag gggaaagccg tacgccaacg ttatttactc agaaacagct 360
tcaacacaca gttgtctgct ttatgatggc atctccaccc aggcacccac catcacctat 420
ctctcgtgcc tgtttatttt cttgcccttt ctgatcataa aaaaacatta agagtttgca 480
aacatgcata ggcatatcaa tatgctcatt tattaatttg ctagcagatc atcttcctac 540
totttacttt atttattgtt tgaaaaatat gtootgoaco tagggagoto gtatacagta 600
ccaatgcate tteattaaat gtgaatttea gaaaggaagt aggaacetat gagagtattt 660
ttcaaaatta attagcggct tctattatgt ttatagcaaa ggccaagggc aaaattggaa 720
cactaatgat ggttggttgc atgagtctgt cgattacttg caagaaatgt gaacctttgt 780
ttctgtgcgt gggcataaaa caaacagctt ctagcctctt ttacggtact tgcacttgca 840
agaaatgtga acteetttte atttetgtat gtggacataa tgccaaagca tecaggettt 900
treatgging treatgeter tacacagine atchecacca grangecore etcatacter 960
atataaacac atcaacagca togcaattag coacaagato acttogggag gcaagtgcga 1020
triceatore geagecacet trittigite tgttgtaagt atacetteee ttaccatett 1080
tatotgttag titaatttgt aattgggaag tattagtgga aagaggatga gatgotatoa 1140
tetatgtact etgeaaatge atetgaegtt atatgggetg etteatataa tttgaattge 1200
tecattette cegacaatat attecaaget atatecetae ttecateaaa aettetett 1260
tttcattcta aaagcatttt agtggcacac aatttttgtc catgagggaa aggaaatctg 1320
ttttggttac tttgcttgag gtgcattctt catatgtcca gttttatgga agtaataaac 1380
ttcagtttgg tcataagatg tcatattaaa gggcaaacat atattcaatg ttcaattcat 1440
cgtaaatgtt ccctttttgt aaaagattgc atactcattt atttgagttg caggtgtatc 1500
                                                                   1515
tagtagttgg aggag
<210> 12
<211> 673
<212> DNA
<213> Zea mays
<400> 12
gatcatccag gtgcaaccgt ataagtccta aagtggtgag gaacacgaaa caaccatgca 60
ttggcatgta aagetecaag aatttgttgt ateettaaca aeteacagaa catcaaccaa 120
aattgcacgt caagggtatt gggtaagaaa caatcaaaca aatcctctct gtgtgcaaag 180
aaacacqqtq agtcatqccq agatcatact catctgatat acatgcttac agctcacaag 240
acattacaaa caactcatat tgcattacaa agatcgtttc atgaaaaata aaataggccg 300
gacaggacaa aaateettga egtgtaaagt aaatttacaa caaaaaaaaa gecatatgte 360
aayctaaatc taattcgttt tacgtagatc aacaacctgt agaaggcaac aaaactgagc 420
cacgcagaag tacagaatga ttccagatga accatcgacg tgctacgtaa agagagtgac 480
gagicatata cattiggicaa gaaaccatga agetgeetae ageegteteg giggicataag 540
aacacaagaa attgtgttaa ttaatcaaag ctataaataa cgctcgcatg cctgtgcact 600
torcoatcac caccactggg tottcagace attagettta totactccag agegeagaag 660
aacccgatcg aca
<210> 13
<211> 454
<212> PRT
<213> Artificial Sequence
<220>
```

<223> synthetic

<400> 13 Met Arg Val Leu Leu Val Ala Leu Ala Leu Leu Ala Leu Ala Ser Ala Thr Ser Ala Lys Tyr Leu Glu Leu Glu Glu Gly Gly Val Ile Met 20 25 Gln Ala Phe Tyr Trp Asp Val Pro Ser Gly Gly Ile Trp Trp Asp Thr 35 40 45 lle Arg Gln Lys Ile Pro Glu Trp Tyr Asp Ala Gly Ile Ser Ala Ile 55 60 Tro Ile Pro Pro Ala Ser Lys Gly Met Ser Gly Gly Tyr Ser Met Gly 70 75 Tyr Asp Prc Tyr Asp Tyr Phe Asp Leu Gly Glu Tyr Tyr Gln Lys Gly 85 90 Thr Val Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Ile Asn Met Ile 100 105 110Asn Thr Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile 115 120 125115 120 Asn His Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Gly Asp 135 140 Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala 150 155 160 Asn Tyr Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly 165 170 Thr Phe Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln 180 190 180 185 Tyr Trp Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser 195 200 Ile Gly Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala 215 210 220 Trp Val Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly 225 230 235 240 Glu Tyr Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Ser 245 250 255 Ser Gly Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Ala 260 265 270 Ala Phe Asp Asn Lys Asn Ile Pro Ala Leu Val Glu Ala Leu Lys Asn 275 280 285 Gly Gly Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val 290 295 300 Ala Ash His Asp Thr Asp Ile Ile Trp Ash Lys Tyr Pro Ala Tyr Ala 305 310 315 320 Phe Ile Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr 325 330 335Glu Glu Trp Leu Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His 345 Asp Asn Leu Ala Gly Gly Ser Thr Ser Ile Val Tyr Tyr Asp Ser Asp 355 360 365 Glu Met Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile 370 380 380 375 Thr Tyr Ile Asn Leu Gly Ser Ser Lys Val Gly Arg Trp Val Tyr Val 385 390 395 400 390 Pro Lys Phe Ala Gly Ala Cys Ile His Glu Tyr Thr Gly Asn Leu Gly 405 410 Gly Trp Val Asp Lys Tyr Val Tyr Ser Ser Gly Trp Val Tyr Leu Glu 420 430 430 Ala Pro Ala Tyr Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp

435 440 445 _ Ser Tyr Cys Gly Val Gly 450 <210> 14 <211> 460 <212> PRT <213> Artificial Sequence <223> synthetic <400> 14 Met Arg Val Leu Leu Val Ala Leu Ala Leu Leu Ala Leu Ala Ser 10 Ala Thr Ser Ala Lys Tyr Leu Glu Leu Glu Glu Gly Gly Val Ile Met 20 25 Gln Ala Phe Tyr Trp Asp Val Pro Ser Gly Gly Ile Trp Trp Asp Thr 35 40 lle Arg Gln Lys Ile Pro Glu Trp Tyr Asp Ala Gly Ile Ser Ala Ile 55 Trp Ile Pro Pro Ala Ser Lys Gly Met Ser Gly Gly Tyr Ser Met Gly 70 75 Tyr Asp Pro Tyr Asp Tyr Phe Asp Leu Gly Glu Tyr Tyr Gln Lys Gly
85 90 95 85 90 95 Thr Val Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Ile Asn Met Ile 105 100 110 Asn Thr Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile 120 115 125 Asn His Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Gly Asp 130 140 135 Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala 145 150 155 160 Asn Tyr Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly 165 170 175 Thr Phe Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln 180 185 Tyr Trp Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser 200 205 Ile Gly Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala 210 225 220 215 Trp Val Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly 225 230 235 240 Glu Tyr Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Ser 245 250 Ser Gly Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Ala 260 265 270 Ala Phe Asp Asn Lys Asn Ile Pro Ala Leu Val Glu Ala Leu Lys Asn 280 285 275 Gly Gly Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val 290 295 300 290 Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala 305 310 315 320 Phe Tle Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr 325 330 335 Glu Glu Trp Leu Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His 350 340 345

```
Asp Asn Leu Ala Gly Gly Ser Thr Ser Ile Val Tyr Tyr Asp Ser Asp
                                           365 I
       355
                         360
Glu Met Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile
                     375
                                      380
Thr Tyr Ile Asn Leu Gly Ser Ser Lys Val Gly Arg Trp Val Tyr Val
                 390
                                    395
385
Pro Lys Phe Ala Gly Ala Cys Ile His Glu Tyr Thr Gly Asn Leu Gly
                        410
            405
Gly Trp Val Asp Lys Tyr Val Tyr Ser Ser Gly Trp Val Tyr Leu Glu
        420
                          425
                                          430
Ala Pro Ala Tyr Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp
    435 440 445
Ser Tyr Cys Gly Val Gly Ser Glu Lys Asp Glu Leu
                     455
<210> 15
<211> 518
<212> PRT
<213> Artificial Sequence
<220>
<223> synthetic
<400> 15
Met Leu Ala Ala Leu Ala Thr Ser Gln Leu Val Ala Thr Arg Ala Gly
                                10
Leu Gly Val Pro Asp Ala Ser Thr Phe Arg Arg Gly Ala Ala Gln Gly
                           25
       20
Leu Arg Gly Ala Arg Ala Ser Ala Ala Ala Asp Thr Leu Ser Met Arg
                       40
      35
Thr Ser Ala Arg Ala Ala Pro Arg His Gln His Gln Gln Ala Arg Arg
             55
                                 60
   50
Gly Ala Arg Phe Pro Ser Leu Val Val Cys Ala Ser Ala Gly Ala Met
                                  75
      7.0
Ala Lys Tyr Leu Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala Phe
                               90
           85
Tyr Trp Asp Val Pro Ser Gly Gly Ile Trp Trp Asp Thr Ile Arg Gln
                           105
                                            110
         100
Lys Ile Pro Glu Trp Tyr Asp Ala Gly Ile Ser Ala Ile Trp Ile Pro
                        120
                                        125
      115
Pro Ala Ser Lys Gly Met Ser Gly Gly Tyr Ser Met Gly Tyr Asp Pro
                  135
                                   140
 130
Tyr Asp Tyr Phe Asp Leu Gly Glu Tyr Tyr Gln Lys Gly Thr Val Glu
                150
                                  155
                                                    160
Thr Arg Phe Gly Ser Lys Gln Glu Leu Ile Asn Met Ile Asn Thr Ala
            165
                              170
                                                 175
His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His Arg
                                              190
          180
                          185
Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Gly Asp Tyr Thr Trp 195 200 205
Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr Leu
210 215 220
 210
           215
Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe Gly
225 230 235
Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp Leu
                                                 255
              245
                              250
Tro Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly Ile
```

```
265
           260
Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val Val
                        280
   275
                                              285
Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp
            295
                                    300
 290
Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Ser Ser Gly Ala 305 \phantom{\bigg|}310\phantom{\bigg|}310\phantom{\bigg|}315\phantom{\bigg|}
Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Ala Ala Phe Asp 325 \phantom{\bigg|}330\phantom{\bigg|}
Asn Lys Asn Ile Pro Ala Leu Val Glu Ala Leu Lys Asn Gly Gly Thr
         340 345
                                              350
Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His
       355
                       360
                                           365
Asp Thr Asp Tie Tie Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Tie Leu 370 380
                     375
The Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu Trp
               390
                            395
Leu Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn Leu
                                410
Ala Gly Gly Ser Thr Ser Ile Val Tyr Tyr Asp Ser Asp Glu Met Ile
                              425
                                                 430
           420
Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr Ile
                          440
                                           445
    435
Asn Leu Gly Ser Ser Lys Val Gly Arg Trp Val Tyr Val Pro Lys Phe
                   455
                                       460
 450
Ala Gly Ala Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp Val
                470
                                     475
465
Asp Lys Tyr Val Tyr Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro Ala
                                490
            485
Tyr Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr Cys
                             505
        500
Gly Val Gly Thr Ser Ile
      515
<210> 16
<211> 820
<212> PRT
<213> Artificial Sequence
<220>
<223> synthetic
<400> 16
Het Leu Ala Ala Leu Ala Thr Ser Gln Leu Val Ala Thr Arg Ala Gly
                                  10
Leu Gly Val Pro Asp Ala Ser Thr Phe Arg Arg Gly Ala Ala Gln Gly
           20
                               25
                                                 30
Leu Arg Gly Ala Arg Ala Ser Ala Ala Ala Asp Thr Leu Ser Met Arg
                         40
                                           45
Thr Ser Ala Arg Ala Ala Pro Arg His Gln His Gln Gln Ala Arg Arg
                  55
                                  60
Gly Ala Arg Phe Pro Ser Leu Val Val Cys Ala Ser Ala Gly Ala Met
                 70
                                     75
Ala Lys Tyr Leu Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala Phe
                                  90
              85
Tyr Trp Asp Val Pro Ser Gly Gly Ile Trp Trp Asp Thr Ile Arg Gln
```

105

Lys	Ile	Pro	Glu	Trp	Tyr	Asp	Ala 120	Gly	Ile	Ser	Ala	Ile 125	Trp_	Ile	Pro
Pro	Ala 130		Lys	Gly	Met	Ser 135		Gly	Tyr	Ser	Met 140	Gl y	Tyr	Asp	Pro
Tyr 145		Tyr	Phe	Asp	Leu 150		Glu	Tyr	Tyr	Gln 155	Lys	Gly	Thr	Val	Glu 160
	Arg	Phe	Gly	Ser 165		Gln	Glu	Leu	Ile 170		Met	Ile	Asn	Thr 175	Ala
His	Ala	Tyr	Gly 180		Lys	Val	Ile	Ala 185	Asp	Ile	Val	Ile	Asn 190	His	Arg
Ala	Gly	Gly 195	Asp	Leu	Glu	Trp	Asn 200	Pro	Phe	Val	Gly	Asp 205	Tyr	Thr	Trp
	210					215					220		Asn		
225					230					235			Thr		Gly 240
				245					250				Tyr	255	
			260					265					11e 270		
		275					280					285	Trp		
	290					295					300		Glu		
305					310					315			Ser		320
Lys	Val			325					330				Ala	335	
	Lys		340					345					Gly 350		
	Val	355					360					365	Ala		
	370					375					380		Phe		
385	-				390					395			Glu		400
		-		405					410				Asp	415	
	•	_	420					425					Glu 430		
		435					440					445	Thr		
	450					455					460		Pro		
465	-				470					475			Gly		480
				485					490				Ala	495	
-			500					505					Ser 510		
-		515					520					525	Arg		
	530					535					540		Ile		
545					550					555			Thr		560
Val	Asn	Gly	Met	Asp 565	Val	Ser	Glu	Trp	Asp 570	Pro	Ser	Arg	Asp	Lys 575	Tyr

Ile Ala Val Lys Tyr Asp Val Ser Thr Ala Val Glu Ala Lys Ala Leu

```
580
                              585
Asn Lys Glu Ala Leu Gln Ala Glu Val Gly Leu Pro Val Asp Arg Asn
     595
                        600
                                            605
 Ile Pro Leu Val Ala Phe Ile Gly Arg Leu Glu Glu Gln Lys Gly Pro
              615
                                        620
Asp Val Met Ala Ala Ala Ile Pro Gln Leu Met Glu Met Val Glu Asp
625
                   630
                                    635
Val Gln Ile Val Leu Leu Gly Thr Gly Lys Lys Phe Glu Arg Met
              645
                                 650
Leu Met Ser Ala Glu Glu Lys Phe Pro Gly Lys Val Arg Ala Val Val
           660
                             665
                                                670
Lys Phe Asn Ala Ala Leu Ala His His Ile Met Ala Gly Ala Asp Val
      675
                         680
                                           685
Leu Ala Val Thr Ser Arg Phe Glu Pro Cys Gly Leu Ile Gln Leu Gln
                                         700
   690
                      695
Gly Met Arg Tyr Gly Thr Pro Cys Ala Cys Ala Ser Thr Gly Gly Leu
705 710 715 720
Val Asp Thr Ile Ile Glu Gly Lys Thr Gly Phe His Met Gly Arg Leu
                               730
             725
Ser Val Asp Cys Asn Val Val Glu Pro Ala Asp Val Lys Lys Val Ala
          740
                       745
                                              750
Thr Thr Leu Gin Arg Ala Ile Lys Val Val Gly Thr Pro Ala Tyr Glu
       755
                  760
                                            765
Glu Met Val Arg Asn Cys Met Ile Gln Asp Leu Ser Trp Lys Gly Pro
  770 775
                                     780
Ala Lys Asn Trp Glu Asn Val Leu Leu Ser Leu Gly Val Ala Gly Gly
785 790
                                    795
Glu Pro Gly Val Glu Gly Glu Glu Ile Ala Pro Leu Ala Lys Glu Asn
            805
                                 810
Val Ala Ala Pro
           820
<210> 17
<211> 19
<212> PRT
<213> Artificial Sequence
<220>
<223> synthetic
<400> 17
Met Arg Val Leu Leu Val Ala Leu Ala Leu Leu Ala Leu Ala Ala Ser
1
                                 10
Ala Thr Ser
<210> 18
<211> 444
<212> PRT
<213> Thermotoga maritima
<400> 18
Met Ala Glu Phe Phe Pro Glu Ile Pro Lys Ile Gln Phe Glu Gly Lys
            5
                               10
                                                  15
Glu Ser Thr Asn Pro Leu Ala Phe Arg Phe Tyr Asp Pro Asn Glu Val
```

```
Ile Asp Gly Lys Pro Leu Lys Asp His Leu Lys Phe Ser Val Ala Phe
                        40
   35
Trp His Thr Phe Val Asn Glu Gly Arg Asp Pro Phe Gly Asp Pro Thr
            55
                                      60
Ala Glu Arg Pro Trp Asn Arg Phe Ser Asp Pro Met Asp Lys Ala Phe 65 70 75 80
Ala Arg Val Asp Ala Leu Phe Glu Phe Cys Glu Lys Leu Asn Ile Glu
85 90 95
Tyr Phe Cys Phe His Asp Arg Asp Ile Ala Pro Glu Gly Lys Thr Leu
                                               110
                      105
          100
Arg Glu Thr Asn Lys Ile Leu Asp Lys Val Val Glu Arg Ile Lys Glu
                                          125
       115
                       120
Arg Met Lys Asp Ser Asn Val Lys Leu Leu Trp Gly Thr Ala Asn Leu
130 135 140
            135
Phe Ser His Pro Arg Tyr Met His Gly Ala Ala Thr Thr Cys Ser Ala
145 150 155 160
Asp Val Phe Ala Tyr Ala Ala Ala Gln Val Lys Lys Ala Leu Glu Ile
            165 170
Thr Lys Glu Leu Gly Gly Glu Gly Tyr Val Phe Trp Gly Gly Arg Glu
                             185
                                               190
         180
Gly Tyr Glu Thr Leu Leu Asn Thr Asp Leu Gly Leu Glu Leu Glu Asn
                                         205
               200
 195
Leu Ala Arg Phe Leu Arg Met Ala Val Glu Tyr Ala Lys Lys Ile Gly 210 215 220
Phe Thr Gly Gln Phe Leu Ile Glu Pro Lys Pro Lys Glu Pro Thr Lys
225 230
                                   235
His Gln Tyr Asp Phe Asp Val Ala Thr Ala Tyr Ala Phe Leu Lys Asn
                   250
             245
His Gly Leu Asp Glu Tyr Phe Lys Phe Asn Ile Glu Ala Asn His Ala 260 265 270
          260
Thr Leu Ala Gly His Thr Phe Gln His Glu Leu Arg Met Ala Arg Ile 275 280 285
Leu Gly Lys Leu Gly Ser Ile Asp Ala Asn Gln Gly Asp Leu Leu Leu
                      295 300
 290
Gly Trp Asp Thr Asp Gln Phe Pro Thr Asn Ile Tyr Asp Thr Thr Leu
                 310 315
Ala Met Tyr Glu Val Ile Lys Ala Gly Gly Phe Thr Lys Gly Gly Leu
325 330 335
Asn Phe Asp Ala Lys Val Arg Arg Ala Ser Tyr Lys Val Glu Asp Leu 340 350
        340
                    345
Phe Ile Gly His Ile Ala Gly Met Asp Thr Phe Ala Leu Gly Phe Lys 355 360 365
Ile Ala Tyr Lys Leu Ala Lys Asp Gly Val Phe Asp Lys Phe Ile Glu
 370 375
Glu Lys Tyr Arg Ser Phe Lys Glu Gly Ile Gly Lys Glu Ile Val Glu
              390
                                    395
Gly Lys Thr Asp Phe Glu Lys Leu Glu Glu Tyr Ile Ile Asp Lys Glu
405 410 415
Asp Ile Glu Leu Pro Ser Gly Lys Gln Glu Tyr Leu Glu Ser Leu Leu
                         425
          420
Asn Ser Tyr Ile Val Lys Thr Ile Ala Glu Leu Arg
                         440
       435
```

<210> 19 <211> 1335 <212> DNA

<213> Thermotoga maritima <400> 19 atggccgagt tcttcccgga gatcccgaag atccagttcg agggcaagga gtccaccaac 60 cogetegeet teegetteta egaceegaac gaggtgateg aeggeaagee geteaaggae 120 cacctcaagt totoogtggo ottotggcac accttogtga acgagggccg cgacccgtto 180 ggcqacccqa ccgccqagcg cccgtggaac cycttotccg acccgatgga caaggcotto 240 geoegegigg acgeeetett egagiteitge gagaagetea acategagia ettetgette 300 cacqueegeg acategoece ggagggeaug acceteegeg agaccuacau gateetegae 360 sagginging agogoatosa ggagogoatg aaggactoca acgigaagot colotigging 420 accgccaacc tettetecca eccgegetac atgeaeggeg eegecaccae etgeteegee 480 gacgtgttcg cctacgccgc cgcccaggtg aagaaggccc tggagatcac caaggagctg 540 ggcggcgagg gctacgtgtt etggggcggc cgcgaggget acgagaccet cetcaacace 600 gaccteggee tggagetgga gaacctegee egetteetee geatggeegt ggagtaegee 660 aagaagatcg gcttcaccgg ccagttcctc atcgagccga agccgaagga gccgaccaag 720 caccagtacg actrogacgt ggccaccgcc tacgccttcc tcaagaacca cggcctcgac 780 gagtactica agiticaacat cgaggccaac cacgccacco tegeoggcca cacettecag 840 cacgagetge geatggeeeg cateetegge aagetegget ceategaege caaccaggge 900 gacetectee teggetggga cacegaceag treeegacea acatetaega caceacete 960 gccatgtacg aggtgatcaa ggccggcggc ttcaccaagg gcggcctcaa cttcgacgcc 1020 aaggtgcgcc gcgcctccta caaggtggag gacctcttca tcggccacat cgccggcatg 1080 gacaccttcg ccctcggctt caagatcgcc tacaagctcg ccaaggacgg cgtgttcgac 1140 aaqttcatcq aggagaagta ccgctccttc aaggagggca tcggcaagga gatcgtggag 1200 ggcaagaccg acttcgagaa gctggaggag tacatcatcg acaaggagga catcgagctg 1260 ccgtccqqca agcaggaqta cctggaqtcc ctcctcaact cctacatcgt gaagaccatc 1320 gccgagetgc getga <210> 20 <211> 444 <212> PRT <213> Thermotoga neapolitana <400> 20 Met Ala Glu Phe Phe Pro Glu Ile Pro Lys Val Gln Phe Glu Gly Lys 10 5 Glu Ser Thr Asn Pro Leu Ala Phe Lys Phe Tyr Asp Pro Glu Glu Ile 25 20 Ile Asp Gly Lys Pro Leu Lys Asp His Leu Lys Phe Ser Val Ala Phe 40 4.5 25 Trp His Thr Phe Val Asn Glu Gly Arg Asp Pro Phe Gly Asp Pro Thr 60 55 50 Ala Asp Arg Pro Trp Asn Arg Tyr Thr Asp Pro Met Asp Lys Ala Phe 75 7.0 Ala Arg Val Asp Ala Leu Phe Glu Phe Cys Glu Lys Leu Asn Ile Glu 90 85 Tyr Phe Cys Phe His Asp Arg Asp Ile Ala Pro Glu Gly Lys Thr Leu 110 105 Arg Glu Thr Asn Lys Ile Leu Asp Lys Val Val Glu Arg Ile Lys Glu 120 125 115 Arg Met Lys Asp Ser Asn Val Lys Leu Leu Trp Gly Thr Ala Asn Leu 140 135 130 Phe Ser His Pro Arg Tyr Met His Gly Ala Ala Thr Thr Cys Ser Ala 150 155 145 Asp Val Phe Ala Tyr Ala Ala Ala Gln Val Lys Lys Ala Leu Glu Ile 170 175 165 Thr Lys Glu Leu Gly Gly Glu Gly Tyr Val Phe Trp Gly Gly Arg Glu 190 185 180 Gly Tyr Glu Thr Leu Leu Asn Thr Asp Leu Gly Phe Glu Leu Glu Asn

Page 21

```
200
                                               205
       195
Leu Ala Arg Phe Leu Arg Met Ala Val Asp Tyr Ala Lys Arg Ile Gly
                                           220
                       215
   210
Phe Thr Gly Gln Phe Leu Ile Glu Pro Lys Pro Lys Glu Pro Thr Lys
                   230
                                       235
His Gln Tyr Asp Phe Asp Val Ala Thr Ala Tyr Ala Phe Leu Lys Ser
                                 250
                                                      255
               245
His Gly Leu Asp Glu Tyr Phe Lys Phe Asn Ile Glu Ala Asn His Ala
                               265
                                                   270
           260
Thr Leu Ala Gly His Thr Phe Gln His Glu Leu Arg Met Ala Arg Ile
                                               285
                           280
       275
Leu Gly Lys Leu Gly Ser Ile Asp Ala Asn Gln Gly Asp Leu Leu
                                           300
                       295
   290
Gly Trp Asp Thr Asp Gln Phe Pro Thr Asn Val Tyr Asp Thr Thr Leu
                                       315
                   310
305
Ala Met Tyr Glu Val Ile Lys Ala Gly Gly Phe Thr Lys Gly Gly Leu
                                                       335
               325
                                   330
Asn Phe Asp Ala Lys Val Arg Arg Ala Ser Tyr Lys Val Glu Asp Leu
                               345
                                                   350
           340
Phe Ile Gly His Ile Ala Gly Met Asp Thr Phe Ala Leu Gly Phe Lys
                                               365
                           360
       355
Val Ala Tyr Lys Leu Val Lys Asp Gly Val Leu Asp Lys Phe Ile Glu
                       375
                                           380
   370
Glu Lys Tyr Arg Ser Phe Arg Glu Gly Ile Gly Arg Asp Ile Val Glu
                                                            400
                                       395
                   390
Gly Lys Val Asp Phe Glu Lys Leu Glu Glu Tyr Ile Ile Asp Lys Glu
                                   410
                                                      415
               405
Thr lie Glu Leu Pro Ser Gly Lys Gin Glu Tyr Leu Glu Ser Leu Ile
                               425
           420
Asn Ser Tyr Ile Val Lys Thr Ile Leu Glu Leu Arg
                            440
<210> 21
<211> 1335
<212> DNA
<213> Thermotoga neapolitana
<400> 21
atggccgagt tettecegga gatecegaag gtgeagtteg agggeaagga gtecaceaac 60
cogeregeet teaagetera egaceeggag gagateateg aeggeaagee geteaaggae 120
caccicaagi iciccgigge ciictggeac accitcgiga acgaggeeg egaccegite 180
ggcgacccga ccgccgaccg cccgtggaac cgctacaccg acccgatgga caaggccttc 240
georgegigg acgeoetett egagitetge gagaagetea acategagia ettetgette 300
cacqaccgcg acatcgcccc ggagggcaag accetccgcg agaccaacaa gatcetcgac 360
aaggtggtgg agcgcatcaa ggagcgcatg aaggactcca acgtgaagct cctctggggc 420
accaccac ectrococa occacactac atacacages esaccac etactecace 480
gacguature cotacgooge egocoaggua aagaaggood uggagatoac caaggagoug 540
ggcggcgagg gctacgtgtt ctggggcggc cgcgagggct acgagaccct cctcaacacc 600
gacctegget tegagetgga gaacctegee egetteetee geatggeegt ggactaegee 660
aagcgcatcg gcttcaccgg ccagttcctc atcgagccga agccgaagga gccgaccaag 720
caccagtacg acttogacgt ggccaccgcc tacgccttcc tcaagtccca cggcctcgac 780
gagtactica agiticaacat egaggecaac caegecaece tegeoggeca caeciticeag 840
cacgagetge geatggeeg cateetegge aagetegget ceategaege caaccaggge 900
gacetectee teggetogga cacegaceag tteeegacea aegtotaega caecaceete 960
gccatgtacg aggtgatcaa ggccggcggc ttcaccaagg gcggcctcaa cttcgacgcc 1020
aaggracec gegeeteeta caaggragag gacetettea teggeeacat egeeggeatg 1080
gacacetteg eceteggett caaggtggee tacaageteg tgaaggaegg egtgetegae 1140
```

```
aagttcatcg aggagaagta cogotootto ogogagggca toggoogga catcgtggag 1200
ggcaaggtgg acttogagaa gotggaggag tacatoatog acaaggagac catogagotg 1260
cogtocogoa agcaggagta cotggagtoc otdatcaact cotacatogt gaagaccate 1320
ctggagetge getga
<210> 22
<211> 28
<212> DNA
<213> Artificial Sequence
<220>
<223> synthetic
<400> 22
                                                                  28
agcgaattca tggcggctct ggccacgt
<210> 23
<211> 29
<212> DNA
<213> Artificial Sequence
<220>
<223> synthetic
<400> 23
                                                                  29
agctaagett cagggegegg ceaegttet
<210> 24
<211> 325
<212> PRT
<213> Artificial Sequence
<220>
<223> synthetic
<400> 24
Met Arg Val Leu Leu Val Ala Leu Ala Leu Leu Ala Leu Ala Ala Ser
                                10
Ala Thr Ser Ala Gly His Trp Tyr Lys His Gln Arg Ala Tyr Gln Phe
                               25
                                                   30
           20
Thr Sly Glu Asp Asp Phe Gly Lys Val Ala Val Val Lys Leu Pro Met
                                               45
                           40
       35
Asp Leu Thr Lys Val Gly Ile Ile Val Arg Leu Asn Glu Trp Gln Ala
                       55
                                            60
Lys Asp Val Ala Lys Asp Arg Phe Ile Glu Ile Lys Asp Gly Lys Ala
                    70
                                        75
65
Glu Val Trp Ile Leu Gln Gly Val Glu Glu Ile Phe Tyr Glu Lys Pro
                                   90
              85
Asp Thr Ser Pro Arg Ile Phe Phe Ala Gln Ala Arg Ser Asn Lys Val
                               105
                                                   110
           100
ile Glu Ala Phe Leu Thr Asn Pro Val Asp Thr Lys Lys Glu Leu
                          120
       115
Phe Lys Val Thr Val Asp Gly Lys Glu Ile Pro Val Ser Arg Val Glu
                       135
                                           140
   130
Lys Ala Asp Pro Thr Asp Ile Asp Val Thr Asn Tyr Val Arg Ile Val
                                     155
                   150
Leu Ser Glu Ser Leu Lys Glu Glu Asp Leu Arg Lys Asp Val Glu Leu
                                   170
               165
```

Ile	Ile	Glu	Gly 180	Tyr	Lys	Pro	Ala	Arg 185	Val	Ile	Met	Met	Glu 190	Ile	Leu
Asp	Asp	Tyr 195	Tyr	Tyr	Asp	Gly	Glu 200		Gly	Ala	Val	Tyr 205		Pro	Glu
•	210	Ile		-	Val	215					220				
225			-		Gly 230					235					240
				245	Asn				250					255	
Leu	-	_	260		Tyr			265					270		
Arg		275			Pro		280					285			
	290				Asn	295					300				
305	-				Lys 310 Asp					315					320
				325	Leu				330					335	
			340		Gly			345					350		
•	Val	355			Pro		360					365			
	370				Lys	375					380				
385	-				390 Gly					395					400
Thr	Arg	Ile	Arg	405 Glu	Val	Lys	Glu		410 Val	Lys	Ala	Leu	His	415 Lys	His
Gly	Ile		420 Val	Ile	Met	Asp		425 Val	Phe	Pro	His		430 Tyr	Gly	Ile
Gly		435 Leu	Ser	Ala	Phe	Asp 4`55	440 Gln	Thr	Val	Pro	Tyr 460	445 Tyr	Phe	Tyr	Arg
	450 Asp	Lys	Thr	Gly	Ala 470		Leu	Asn	Glu	Ser 475		Cys	Gly	Asn	Val 480
465 Ile	Ala	Ser	Glu	Arg 485	Pro	Met	Met	Arg	Lys 490		Ile	Val	Asp	Thr 495	
Thr	Tyr	Trp	Val 500		Glu	Tyr	Hıs	Ile 505		Gly	Phe	Arg	Phe 510	Asp	Gln
		515			Lys		520					525			
	530				Thr	535					540				
545	-				Arg 550					555					560
				565	Asp				570					575	
			580		Lys			585					590		
	•	595			Gly		600					605			
	610	-			Ala	615					620				
A1a 625	cys	HIS	АЅР	ASN	630	101	րգդ	тþ	vsh	635	USII	: A I	Pen	Y.O	640

```
Lys Ala Asp Lys Lys Clu Trp Thr Glu Glu Glu Leu Lys Asn Ala
                                   650
               645
Gln Lys Leu Ala Gly Ala Ile Leu Leu Thr Ser Gln Gly Val Pro Phe
                               665
                                                  670
           660
Leu His Gly Gly Gln Asp Phe Cys Arg Thr Thr Asn Phe Asn Asp Asn
                          680
       675
Ser Tyr Asn Ala Pro Ile Ser Ile Asn Gly Phe Asp Tyr Glu Arg Lys
    690
                       695
Leu Gla Phe Ile Asp Val Phe Asn Tyr His Lys Gly Leu Ile Lys Leu
                                       715
                   710
Arg Lys Glu His Pro Ala Phe Arg Leu Lys Asn Ala Glu Glu Ile Lys
                                  730
               725
Lys His Leu Glu Phe Leu Pro Gly Gly Arg Arg Ile Val Ala Phe Met
                               745
                                                   750
           740
Leu Lys Asp His Ala Gly Gly Asp Pro Trp Lys Asp Ile Val Val Ile
                                               765
                           760
       755
Tyr Asn Gly Asn Leu Glu Lys Thr Thr Tyr Lys Leu Pro Glu Gly Lys
                                           780
                      775
Trp Asn Val Val Val Asn Ser Gln Lys Ala Gly Thr Glu Val Ile Glu
                                       795
                   790
785
Thr Val Glu Gly Thr Ile Glu Leu Asp Pro Leu Ser Ala Tyr Val Leu
                        810
              805
Tyr Arg Glu Ser Glu Lys Asp Glu Leu
           820
<210> 25
<211> 2478
<212> DNA
<213> Artificial Sequence
<220>
<223> synthetic
<400> 25
atgagggtgt tgctcgttgc cctcgctctc ctggctctcg ctgcgagcgc caccagcgct 60
ggocactiggt acaagcacca gogogootac cagttoacog gogaggacga ottogggaag 120
gtggccgtgg tgaagctccc gatggacctc accaaggtgg gcatcatcgt gcgcctcaac 180
gagtggcagg cgaaggacgt ggccaaggac cgcttcatcg agatcaagga cggcaaggcc 240
gaggtgtggga tactccaggg cgtggaggag atcttctacg agaagccgga cacctccccg 300
egeatettet tegeceagge eegeteeaac aaggtgateg aggeetteet caccaaceeg 360
gtggacacca agaagaagga gctgttcaag gtgaccgtcg acggcaagga gatcccggtg 420
tecegegtgg agaaggeega eeegaeegae ategaegtga eeaactaegt gegeategtg 480
ctctccgagt ccctcaagga ggaggacctc cgcaaggacg tggagctgat catcgagggc 540
tacaageegg eccgegtgat catgatggag atcetegacg actactacta egacggegag 600
ctgggggcgg tgtactcccc ggagaagacc atcttccgcg tgtggtcccc ggtgtccaag 660
tgggtgaagg tgctcctctt caagaacggc gaggacaccg agccgtacca ggtggtgaac 720
atggagtaca agggcaacgg cgtgtgggag gccgtggtgg agggcgacct cgacggcgtg 780
ttotacotot accagotoga gaactacogo aagatoogoa coaccotoga cocotactoo 840
aaggoogtgt acgocaacaa coaggagtot goagtggtga acctogooog caccaaccog 900
gagggotggg agaacgaceg eggcocgaag ategaggget acgaggacge catcatetae 960
gagatocaca togoogacat cacoggootg gagaactoog gogtgaagaa caagggooto 1020
taccteggee teacegagga gaacaceaag geecegggeg gegtgaceae eggeetetee 1080
cacctcgtgg agetgggcgt gacccacgtg cacatcctcc cgttcttcga cttctacacc 1140
ggcgacgage tggacaagga ettegagaag tactacaaet ggggetaega ecegtaeete 1200
ttcatggtgc cggagggccg ctactccacc gacccgaaga acccgcacac ccgaattcgc 1260
gaçqıqaaqq agarqqıçaa ggccctccac aagcacqqca tcqqcqtqat catqqacatq 1320
gugitocogo acacetacyg categorgag etgicogect tegaceagae egigeogiae 1380
```

```
tacttctacc qcatcqacaa gaccqqcqcc tacctcaacq aqtccqqctq cqqcaacqtq 1440
ategecteeg agegeeegat gatgegeaag ticategtgg acacegtgae etactgggtg 1500
aaqqaqtacc acategaegg etteegette gaccagatgg geeteatega caagaagaee 1560
atgctggagg tggagcgcgc cctccacaag atcgacccga ccatcatcct ctacggcgag 1620
ceqtqqqqqq qctqqqqqq cccqatccqc ttcqqcaaqt ccqacqtqqc cqqcacccac 1680
gragecact teaacgacga attecacgae accatecaca getecatatt caaccatee 1740
grgaaggget tegrgarggg eggetaegge aaggagaeea agateaageg eggegrggtg 1800
gactocatea actacgaegg caageteate aagteetteg eeetegaeee ggaggagaee 1860
atcaactacg eegectgeea egacaaceae accetetggg acaagaacta cetegeegee 1920
aaggoogaca agaagaagga gtggacogag gaggagotga agaacgooca gaagotogoo 1980
ggcgccatcc tecteactag teagggegtg cegtteetee aeggeggeea ggaettetge 2040
cgcaccacca acttcaacga caactcctac aacgccccga tctccatcaa cggcttcgac 2100
tacqaqcqca agctccaqtt catcqacqtq ttcaactacc acaagggcct catcaagctc 2160
cgcaaggage accoggeett cegeeteaag aacgeegagg agateaagaa geacetggag 2220
ttcctcccgg gcgggcgccg catcgtggcc ttcatgctca aggaccacgc cggcggcgac 2280
ccgtggaagg acatcgtggt gatctacaac ggcaacctgg agaagaccac ctacaagctc 2340
ccggagggca agtggaacgt ggtggtgaac tcccagaagg ccggcaccga ggtgatcgag 2400
accqtqqaqq qcaccatcqa qctqqacccq ctctccqcct acqtqctcta ccqcqaqtcc 2460
gagaaggacg agctgtga
<210> 26
<211> 718
<212> PRT
<213> Artificial Sequence
<220>
<223> synthetic
<400> 26
Met Arg Val Leu Leu Val Ala Leu Ala Leu Ala Leu Ala Ser
                            10
Ala Thr Ser Met Glu Thr lle Lys Ile Tyr Glu Asn Lys Gly Val Tyr
                               25
Lys Val Val Ile Gly Glu Pro Phe Pro Pro Ile Glu Phe Pro Leu Glu
                           40
Gln Lys Ile Ser Ser Asn Lys Ser Leu Ser Glu Leu Gly Leu Thr Ile
                       55
                                           60
Val Gln Gln Gly Asn Lys Val Ile Val Glu Lys Ser Leu Asp Leu Lys
                                       75
Glu His Ile Ile Gly Leu Gly Glu Lys Ala Phe Glu Leu Asp Arg Lys
                                   90
               85
Arg Lys Arg Tyr Val Met Tyr Asn Val Asp Ala Gly Ala Tyr Lys Lys
                               105
                                                   110
           100
Tyr Gln Asp Pro Leu Tyr Val Ser Ile Pro Leu Phe Ile Ser Val Lys
                          120
       115
Asp Gly Val Ala Thr Gly Tyr Phe Phe Asn Ser Ala Ser Lys Val Ile
                                           140
                       135
Phe Asp Val Gly Leu Glu Glu Tyr Asp Lys Val Ile Val Thr Ile Pro
                                       155
                   150
Glu Asp Ser Val Glu Phe Tyr Val Ile Glu Gly Pro Arg Ile Glu Asp
                                                    175
                                   170
               165
Val Leu Glu Lys Tyr Thr Glu Leu Thr Gly Lys Pro Phe Leu Pro Pro
                                                  190
           180
                               185
Met Trp Ala Phe Gly Tyr Met Ile Ser Arg Tyr Ser Tyr Tyr Pro Gln
                           200
                                               205
Asp Lys Val Val Glu Leu Val Asp Ile Met Gln Lys Glu Gly Phe Arg
                       215
                                           220
   210
Val Ala Gly Val Phe Leu Asp Ile His Tyr Met Asp Ser Tyr Lys Leu
```

225					230					235					240
	Thr	Trp	His	Pro 245	Tyr		Phe	Pro	Gl u 250	Pro	Lys	Lys	Leu	11e 255	
Glu	Leu	His	Lys 260	Arg	Asn	Val	Lys	Leu 265			Ile	Val	Asp 270		Gly
Ile	Arg	Val 275			Asn	Tyr	Ser 280		Phe	Leu	Ser	Gly 285		Gly	Lys
Phe	Cys 290		lle	Glu	Ser	Gly 295		Leu	Phe	Val	Gly 300		Met	Trp	Pro
Gly 305		Thr	Val	Tyr	Pro 310		Phe	Phe	Arg	Glu 315		Thr	Arg	Glu	Trp 320
	Ala	Gly	Leu	11e 325	Ser	Glu	Trp	Leu	Ser 330		Gly	Val	Asp	Gly 335	
Trp	Leu	Asp	Met 340		Glu	Pro	Thr	Asp 345	Phe	Ser	Arg	Ala	Ile 350	Glu	Ile
Arg	Asp	Val 355	Leu	Ser	Ser	Leu	Pro 360	Val	Gln	Phe	Arg	Asp 365	Asp	Arg	Leu
Val	370				Asp	375				-	380	•	•	_	_
Val 385					Val 390					395					400
				405	Phe				410					415	
		•	420	•	Tyr		-	425		•	-		430	Ile	•
	_	435			Pro		440					445			
	450				Ser	455					460				
465	-	_			Gly 470					475					480
	Leu			485			Ala		490					495	
Arg			500		Thr			505					510		
	Asp	515	•	-	Glu	-	520					525			
-	530				Ile Arg	535					540				
545					550 Glu					555					560
				565	Ser				570					575	
•			580		Asn	-		585					590		
-		595			Thr		600					605			
	610					615					620				
625					Glu 630					635					640
				645	Asp				650					655	
	-		660		Glu			665					670		
	-	675			Lys		680					685			
Val	Glu	Lys	Thr	Met	Gln	Asn	Thr	Tyr	Val	Ala	Lys	Ile	Asn	Gln	Lys

```
695
Ile Arg Gly Lys Ile Asn Leu Glu Ser Glu Lys Asp Glu Leu
                  710
<210> 27
<211> 712
<212> PRT
<213> Artificial Sequence
<220>
<223> synthetic
<400> 27
Met Arg Val Leu Leu Val Ala Leu Ala Leu Leu Ala Leu Ala Ala Ser
                                10
Ala Thr Ser Met Glu Thr Ile Lys Ile Tyr Glu Asn Lys Gly Val Tyr
                                              3.0
                           25
       20
Lys Val Val Ile Gly Glu Pro Phe Pro Pro Ile Glu Phe Pro Leu Glu
                        40
                                            45
 35
Gln Lys Ile Ser Ser Asn Lys Ser Leu Ser Glu Leu Gly Leu Thr Ile
                   55
 50
Val Gln Gln Gly Asn Lys Val Ile Val Glu Lys Ser Leu Asp Leu Lys 65 70 75
Glu His Ile Ile Gly Leu Gly Glu Lys Ala Phe Glu Leu Asp Arg Lys
85 90 95
                                90
            8.5
Arg Lys Arg Tyr Val Met Tyr Asn Val Asp Ala Gly Ala Tyr Lys Lys 100 105 110
Tyr Gin Asp Pro Leu Tyr Val Ser Ile Pro Leu Phe Ile Ser Val Lys
                      120
                                         125
    115
Asp Gly Val Ala Thr Gly Tyr Phe Phe Asn Ser Ala Ser Lys Val Ile
130 135 140
Phe Asp Val Gly Leu Giu Glu Tyr Asp Lys Val Ile Val Thr Ile Pro
145 150 155 160
       150
Glu Asp Ser Val Glu Phe Tyr Val Ile Glu Gly Pro Arg Ile Glu Asp
          165
                              170
Val Leu Glu Lys Tyr Thr Glu Leu Thr Gly Lys Pro Phe Leu Pro Pro
                                      190
                  185
         180
Met Trp Ala Phe Gly Tyr Met Ile Ser Arg Tyr Ser Tyr Tyr Pro Gln
200 205
                          200
    195
Asp Lys Val Val Glu Leu Val Asp Ile Met Gln Lys Glu Gly Phe Arg 210 215 220
Val Ala Gly Val Phe Leu Asp Ile His Tyr Met Asp Ser Tyr Lys Leu 235 240
Phe Thr Trp His Pro Tyr Arg Phe Pro Glu Pro Lys Lys Leu Ile Asp
245 250 255
           245
Glu Leu His Lys Arg Asn Val Lys Leu Ile Thr Ile Val Asp His Gly
                           265
                                            270
          260
lle Arg Val Asp Gln Asn Tyr Ser Pro Phe Leu Ser Gly Met Gly Lys
      275 280
                                   285
Phe Cys Glu Ile Glu Ser Gly Glu Leu Phe Val Gly Lys Met Trp Pro 290 295 300
Gly Thr Thr Val Tyr Pro Asp Phe Phe Arg Glu Asp Thr Arg Glu Trp
               310 315
Trp Ala Gly Leu Ile Ser Glu Trp Leu Ser Gln Gly Val Asp Gly Ile
                          330
           325
                                                   335
Trp Leu Asp Met Asn Glu Pro Thr Asp Phe Ser Arg Ala Ile Glu Ile
                            345
           340
```

```
Arg Asp Val Leu Ser Ser Leu Pro Val Gln Phe Arg Asp Asp Arg Leu
      355
                        360
Val Thr Thr Phe Pro Asp Asn Val Val His Tyr Leu Arg Gly Lys Arg
                                    380
                 375
Val Lys His Glu Lys Val Arg Asn Ala Tyr Pro Leu Tyr Glu Ala Met
                 390
                         395
Ala Thr Phe Lys Gly Phe Arg Thr Ser His Arg Asn Glu Ile Phe Ile
                                              415
                           410
            405
Leu Ser Arg Ala Gly Tyr Ala Gly Ile Gln Arg Tyr Ala Phe Ile Trp
         420
                           425
                                       430
Thr Gly Asp Asn Thr Pro Ser Trp Asp Asp Leu Lys Leu Gln Leu Gln 445
Leu Val Leu Gly Leu Ser Ile Ser Gly Val Pro Phe Val Gly Cys Asp
                 455
                                   460
  450
Ile Gly Gly Phe Gln Gly Arg Asn Phe Ala Glu Ile Asp Asn Ser Met
              470
                                475
465
Asp Leu Leu Val Lys Tyr Tyr Ala Leu Ala Leu Phe Phe Pro Phe Tyr
                                         495
                              490
             485
Arg Ser His Lys Ala Thr Asp Gly Ile Asp Thr Glu Pro Val Phe Leu
                                     510
                          505
          500
Pro Asp Tyr Tyr Lys Glu Lys Val Lys Glu Ile Val Glu Leu Arg Tyr
                                  525
                520
      515
Lys Phe Leu Pro Tyr Ile Tyr Ser Leu Ala Leu Glu Ala Ser Glu Lys
                                    540
  530
                  535
Gly His Pro Val Ile Arg Pro Leu Phe Tyr Glu Phe Gln Asp Asp Asp
                                   555
              550
Asp Met Tyr Arg Ile Glu Asp Glu Tyr Met Val Gly Lys Tyr Leu Leu
                     570
             565
Tyr Ala Pro Ile Val Ser Lys Glu Glu Ser Arg Leu Val Thr Leu Pro
        580
                  585
Arg Gly Lys Trp Tyr Asn Tyr Trp Asn Gly Glu Ile Ile Asn Gly Lys
                                      605
     595
                      600
Ser Val Val Lys Ser Thr His Glu Leu Pro Ile Tyr Leu Arg Glu Gly
                                     620
                   615
  610
Ser Ile Ile Pro Leu Glu Gly Asp Glu Leu Ile Val Tyr Gly Glu Thr
                                 635
           630
Ser Phe Lys Arg Tyr Asp Asn Ala Glu Ile Thr Ser Ser Ser Asn Glu 655
lle Lys Phe Ser Arg Glu Ile Tyr Val Ser Lys Leu Thr Ile Thr Ser
                                    670
       660
                           665
Glu Lys Pro Val Ser Lys Ile Ile Val Asp Asp Ser Lys Glu Ile Gln
                                  685
      675 680
Val Glu Lys Thr Met Gln Asn Thr Tyr Val Ala Lys Ile Asn Gln Lys
          695
                                      700
The Arg Gly Lys The Asn Leu Glu
                 710
<210> 28
<211> 469
<212> PRT
<213> Artificial Sequence
<220>
<223> synthetic
<400> 28
```

Met Arg Val Leu Leu Val Ala Leu Ala Leu Leu Ala Leu Ala Ala Ser

1				5					10					15	
			20					25	Glu				30	Gln	
		35					40					45		Asp	
	50					55					60			Phe	
65					70					75				Phe	80
				85					90					Met 95	
-			100					105					110	Lys	
		115					120					125		Glu	
	130					135					140			Glu	
145					150					155				Gly	160
				Ser 165					170					Thr 175	
-			180					185					190	Lys	
		195					200					205		Trp	
-	210					215					220			Leu	
225					230					235				Ala	240
-				245					250					255 Ala	
		-	260					265					270	Glu	
		275					280					285		Arg	
	290					295					300			Gly	
305					310					315				Tyr	320
				325					330					335 Thr	
			340					345					350	Lys	
•	-	355					360					365		Ala	
	370					375					380			Asp	
385					390					395					400 Glu
				405					410					415 Ile	
			420					425					430	Leu	
•	-	435					440					445		Arg	
	450	Asp			. , .	455		-,-			460			,	
U L U	-10	٠.٠													

465

```
<210> 29
<211> 469
<212> PRT
<213> Artificial Sequence
<210>
<223> synthetic
<400> 29
Met Arg Val Leu Leu Val Ala Leu Ala Leu Leu Ala Leu Ala Ala Ser
1 5 10 15
Ala Thr Ser Met Ala Glu Phe Phe Pro Glu Ile Pro Lys Val Gln Phe
                        25
 20
Glu Gly Lys Glu Ser Thr Asn Pro Leu Ala Phe Lys Phe Tyr Asp Pro
                                        45
                     40
 35
Glu Glu Ile Ile Asp Gly Lys Pro Leu Lys Asp His Leu Lys Phe Ser
                                    60
                 55
Val Ala Phe Trp His Thr Phe Val Asn Glu Gly Arg Asp Pro Phe Gly
       70
                             75
Asp Pro Thr Ala Asp Arg Pro Trp Asn Arg Tyr Thr Asp Pro Met Asp 85 90 95
Lys Ala Phe Ala Arg Val Asp Ala Leu Phe Glu Phe Cys Glu Lys Leu
       100 105
                                          110
Asn Ile Glu Tyr Phe Cys Phe His Asp Arg Asp Ile Ala Pro Glu Gly
    115 120
                                      125
Lys Thr Leu Arg Glu Thr Asn Lys Ile Leu Asp Lys Val Val Glu Arg
 130 135 140
Ile Lys Glu Arg Met Lys Asp Ser Asn Val Lys Leu Leu Trp Gly Thr
145 150 155 160
Ala Asn Leu Phe Ser His Pro Arg Tyr Met His Gly Ala Ala Thr Thr
165 170 175
Cys Ser Ala Asp Val Phe Ala Tyr Ala Ala Ala Gln Val Lys Lys Ala
       180 185 190
Leu Glu Ile Thr Lys Glu Leu Gly Gly Glu Gly Tyr Val Phe Trp Gly
                                205
 195
                     200
Gly Arg Glu Gly Tyr Glu Thr Leu Leu Asn Thr Asp Leu Gly Phe Glu
210 215 220
Leu Glu Asn Leu Ala Arg Phe Leu Arg Met Ala Val Asp Tyr Ala Lys
225 230 235 240
Arg Ile Gly Phe Thr Gly Gln Phe Leu Ile Glu Pro Lys Pro Lys Glu
          245 250
                                             255
Pro Thr Lys His Gln Tyr Asp Phe Asp Val Ala Thr Ala Tyr Ala Phe 260 270
         260 265
Leu Lys Ser His Gly Leu Asp Glu Tyr Phe Lys Phe Asn Ile Glu Ala
275 280 285
Asn His Ala Thr Leu Ala Gly His Thr Phe Gln His Glu Leu Arg Met
           295 300
Ala Arg Ile Leu Gly Lys Leu Gly Ser Ile Asp Ala Asn Gln Gly Asp 305 310 315
Leu Leu Gly Trp Asp Thr Asp Gln Phe Pro Thr Asn Val Tyr Asp
                                                335
            325 330
Thr Thr Leu Ala Met Tyr Glu Val Ile Lys Ala Gly Gly Phe Thr Lys
                     345 350
        340
Gly Gly Leu Asn Phe Asp Ala Lys Val Arg Arg Ala Ser Tyr Lys Val
                        360
```

```
Glu Asp Leu Phe Ile Gly His Ile Ala Gly Met Asp Thr Phe Ala Leu
                      375
 370
Gly Phe Lys Val Ala Tyr Lys Leu Val Lys Asp Gly Val Leu Asp Lys
                                   395
               390
Phe Ile Glu Glu Lys Tyr Arg Ser Phe Arg Glu Gly Ile Gly Arg Asp 405 \hspace{1.5cm} 410 \hspace{1.5cm} 415
Ile Val Glu Gly Lys Val Asp Phe Glu Lys Leu Glu Glu Tyr Ile Ile
                 425
                                              430
          420
Asp Lys Glu Thr Ile Glu Leu Pro Ser Gly Lys Gln Glu Tyr Leu Glu
               440
                                      445
    435
Ser Leu Ile Asn Ser Tyr Ile Val Lys Thr Ile Leu Glu Leu Arg Ser
                   455
Glu Lys Asp Glu Leu
465
<210> 30
<211> 463
<212> PRT
<213> Artificial Sequence
<220>
<223> synthetic
<400> 30
Met Arg Val Leu Leu Val Ala Leu Ala Leu Leu Ala Leu Ala Ala Ser
                                10
Ala Thr Ser Met Ala Glu Phe Phe Pro Glu Ile Pro Lys Val Gln Phe
                          25
                                             30
       20
Glu Gly Lys Glu Ser Thr Asn Pro Leu Ala Phe Lys Phe Tyr Asp Pro
                      40 45
    35
Glu Glu Ile Ile Asp Gly Lys Pro Leu Lys Asp His Leu Lys Phe Ser
                     5.5
                                       60
Val Ala Phe Trp His Thr Phe Val Asn Glu Gly Arg Asp Pro Phe Gly
                  70
                                     75
Asp Pro Thr Ala Asp Arg Pro Trp Asn Arg Tyr Thr Asp Pro Met Asp 85 90 95
Lys Ala Phe Ala Arg Val Asp Ala Leu Phe Glu Phe Cys Glu Lys Leu
100 105
        100 105
Asn Ile Glu Tyr Phe Cys Phe His Asp Arg Asp Ile Ala Pro Glu Gly
                    120
                                          125
 115
Lys Thr Leu Arg Glu Thr Asn Lys Ile Leu Asp Lys Val Val Glu Arg
                                     140
                   135
Ile Lys Glu Arg Met Lys Asp Ser Asn Val Lys Leu Leu Trp Gly Thr
                             155
               150
145
Ala Asn Leu Phe Ser His Pro Arg Tyr Met His Gly Ala Ala Thr Thr 165 170 175
Cys Ser Ala Asp Val Phe Ala Tyr Ala Ala Ala Gln Val Lys Lys Ala
180 185 190
Leu Glu Ile Thr Lys Glu Leu Gly Gly Glu Gly Tyr Val Phe Trp Gly
                                          205
                         200
Gly Arg Glu Gly Tyr Glu Thr Leu Leu Asn Thr Asp Leu Gly Phe Glu
                   215
 210
Leu Glu Asn Leu Ala Arg Phe Leu Arg Met Ala Val Asp Tyr Ala Lys
           230
                            235 ·
Arg Ile Gly Phe Thr Gly Gln Phe Leu Ile Glu Pro Lys Pro Lys Glu
             245
                             250
                                                    255
Pro Thr Lys His Gln Tyr Asp Phe Asp Val Ala Thr Ala Tyr Ala Phe
```

```
265
                                            270
         260
Leu Lys Ser His Gly Leu Asp Glu Tyr Phe Lys Phe Asn Ile Glu Ala
  275
                280
Asn His Ala Thr Leu Ala Gly His Thr Phe Gln His Glu Leu Arg Met
 290 295
                                  300
Ala Arg Ile Leu Gly Lys Leu Gly Ser Ile Asp Ala Asn Gln Gly Asp
305 310 315
Leu Leu Cly Trp Asp Thr Asp Gln Phe Pro Thr Asn Val Tyr Asp
                             330
             325
Thr Thr Leu Ala Met Tyr Glu Val Ile Lys Ala Gly Gly Phe Thr Lys
         340 345
Gly Gly Leu Asn Phe Asp Ala Lys Val Arg Arg Ala Ser Tyr Lys Val 355 360 365
Glu Asp Leu Phe Ile Gly His Ile Ala Gly Met Asp Thr Phe Ala Leu
                                 380
          375
Gly Phe Lys Val Ala Tyr Lys Leu Val Lys Asp Gly Val Leu Asp Lys
                                 395
              390
Phe Ile Glu Glu Lys Tyr Arg Ser Phe Arg Glu Gly Ile Gly Arg Asp
                                            415
            405
                            410
lle Val Glu Gly Lys Val Asp Phe Glu Lys Leu Glu Glu Tyr Ile Ile
                                    430
                  425
        420
Asp Lys Glu Thr Ile Glu Leu Pro Ser Gly Lys Gln Glu Tyr Leu Glu
435 440
                                  445
Ser Leu Ile Asn Ser Tyr Ile Val Lys Thr Ile Leu Glu Leu Arg
                 455
<210> 31
<211> 25
<212> PRT
<213> Artificial Sequence
<220>
<223> synthetic
<400> 31
Met Gly Lys Asn Gly Asn Leu Cys Cys Phe Ser Leu Leu Leu Leu
Leu Ala Gly Leu Ala Ser Gly His Gln
 20
<210> 32
<211> 30
<212> PRT
<213> Artificial Sequence
<220>
<223> synthetic
<400> 32
Met Gly Phe Val Leu Phe Ser Gln Leu Pro Ser Phe Leu Leu Val Ser
                           10
) 5
Thr Leu Leu Phe Leu Val Ile Ser His Ser Cys Arg Ala
          20
```

<210> 33

<211> 460 <212> PRT <213> Artificial Sequence <220> <223> synthetic <400> 33 Met Arg Val Leu Leu Val Ala Leu Ala Leu Leu Ala Leu Ala Ala Ser 10 15 Ala Thr Ser Ala Lys Tyr Leu Glu Leu Glu Glu Glu Gly Val Ile Met 25 20 Gln Ala Phe Tyr Trp Asp Val Pro Ser Gly Gly Ile Trp Trp Asp Thr 40 4.5 lle Arg Gln Lys Ile Pro Glu Trp Tyr Asp Ala Gly Ile Ser Ala Ile 60 50 55 Trp lie Pro Pro Ala Ser Lys Gly Met Ser Gly Gly Tyr Ser Met Gly 55 70 75 Tyr Asp Pro Tyr Asp Tyr Phe Asp Leu Gly Glu Tyr Tyr Gln Lys Gly 85 90 Thr Val Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Ile Asn Met Ile 100 105 110 105 100 Asn Thr Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile 120 115 Asn His Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Gly Asp 130 135 140 Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala 145 150 155 Asn Tyr Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly 165 170 Thr Phe Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln 180 Tyr Trp Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser 195 200 205 Ile Gly Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala 210 215 220 Trp Val Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly 225 230 240 225 230 Glu Tyr Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Ser 245 250 255 Ser Gly Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Ala 260 270 Ala Phe Asp Asn Lys Asn Ile Pro Ala Leu Val Glu Ala Leu Lys Asn 275 280 285 Gly Gly Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val 295 300 Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala 305 310 315 310 305 Phe Ile Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr 325 330 335 Glu Glu Trp Leu Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His 340 345 350 345 350 340 Asp Asn Leu Ala Gly Gly Ser Thr Ser Ile Val Tyr Tyr Asp Ser Asp 360 355 Glu Met Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile 370 375 380 Thr Tyr Ile Asn Leu Gly Ser Ser Lys Val Gly Arg Trp Val Tyr Val 395

Fro Lys Phe Ala Gly Ala Cys Ile His Glu Tyr Thr Gly Asn Leu Gly 405 410 415 Gly Trp Val Asp Lys Tyr Val Tyr Ser Ser Gly Trp Val Tyr Leu Glu 420 425 430 Ala Pro Ala Tyr Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp 4 4 0 435 Ser Tyr Cys Gly Val Gly Ser Glu Lys Asp Glu Leu 455 <210> 34 <211> 825 <212> PRT <213> Artificial Sequence <220> <223> synthetic <400> 34 Met Arg Val Leu Leu Val Ala Leu Ala Leu Leu Ala Leu Ala Ala Ser 5 10 Ala Thr Ser Ala Gly His Trp Tyr Lys His Gln Arg Ala Tyr Gln Phe 20 25 Thr Gly Glu Asp Asp Phe Gly Lys Val Ala Val Val Lys Leu Pro Met 40 Asp Leu Thr Lys Val Gly Ile Ile Val Arg Leu Asn Glu Trp Gln Ala 55 Lys Asp Val Ala Lys Asp Arg Phe Ile Glu Ile Lys Asp Gly Lys Ala -70 75 65 Glu Val Trp Ile Leu Gln Gly Val Glu Glu Ile Phe Tyr Glu Lys Pro 85 90 Asp Thr Ser Pro Arg Ile Phe Phe Ala Gln Ala Arg Ser Asn Lys Val 100 110 105 Ile Glu Ala Phe Leu Thr Asn Pro Val Asp Thr Lys Lys Lys Glu Leu 120 115 125 Phe Lys Val Thr Val Asp Gly Lys Glu Ile Pro Val Ser Arg Val Glu 130 135 140 Lys Ala Asp Pro Thr Asp Ile Asp Val Thr Asn Tyr Val ${\sf Arg}$ Ile Val 145 150 155 Leu Ser Glu Ser Leu Lys Glu Glu Asp Leu Arg Lys Asp Val Glu Leu 170 165 175 Ile Ile Glu Gly Tyr Lys Pro Ala Arg Val Ile Met Met Glu Ile Leu 180 185 190 Asp Asp Tyr Tyr Tyr Asp Gly Glu Leu Gly Ala Val Tyr Ser Pro Glu 200 Lys Thr Ile Phe Arg Val Trp Ser Pro Val Ser Lys Trp Val Lys Val 210 220 Leu Leu Phe Lys Asn Gly Glu Asp Thr Glu Pro Tyr Gln Val Val Asn 230 235 Met Glu Tyr Lys Gly Asn Gly Val Trp Glu Ala Val Val Glu Gly Asp 245 250 255 Leu Asp Gly Val Phe Tyr Leu Tyr Gln Leu Glu Asn Tyr Gly Lys Ile 260 265 270 Arg Thr Thr Val Asp Pro Tyr Ser Lys Ala Val Tyr Ala Asn Asn Gln 275 280 285 280 Glu Ser Ala Val Val Asn Leu Ala Arg Thr Asn Pro Glu Gly Trp Glu 295 300 Asn Asp Arg Gly Pro Lys Ile Glu Gly Tyr Glu Asp Ala Ile Ile Tyr

305					310					315					320
305 Glu	Ile	His	Ile	Ala 325	Asp	Ile	Thr	Gly	Leu 330		Asn	Ser	Gly	Val 335	
Asn	Lys	Gly	Leu 340	Tyr	Leu	Gly	Leu	Thr 345		Glu	Asn	Thr	Lys 350	Ala	Pro
•	•	355	Thr				360					365		Val	
	370					375					380			Glu	
Asp 385	Lys				390					395				Tyr	400
				405					410					Pro 415	
	-		420					425					430	Lys	
		4.35		Ile			440					445	•	Gly	
-	450					455					460			Tyr	
465					470					475				Asn	480
				485					490					495 Asp	
	•	-	500					505					510	Ala	
		515					520					525		Gly	
	รรัก		-			535					540			Thr	
505					550					555				Ser	טטכ
				565					570					575 Lys	
			580					585					590	Gly	
		595					600					605		Tyr	
	610					615					620			Ala	Ala
625					630				Glu	635				Asn	640
Gln	Lys	Leu		645 Gly	Ala	Ile	Leu		650 Thr	Ser	Gln	Gly	Val	655 Pro	Phe
Leu	His		660 Gly	Gln	Asp	Phe	Cys	665 Arg	Thr	Thr	Asn	Phe	670 Asn	Asp	Asn
Ser		675 Asn	Ala	Pro	Ile	Ser	680 Ile	Asn	Gly	Phe	Asp 700	685 Tyr	Glu	Arg	Lys
	690 Gln	Phe	Ile	Asp	Val 710	695 Phe	Asn	Tyr	His	Lys 715		Leu	Ile	Lys	Leu 720
705 Arg	Lys	Glu	His	Pro 725	Ala	Phe	Arg	Leu	Lys 730		Ala	Glu	Glu	Ile 735	
Lys	His	Leu	Glu 740	Phe	Leu	Pro	Gly	Gly 745		Arg	Ile	Val	Ala 750	Phe	Met
		755	Hıs				760	Pro				765	Val	Val	
Tyr	Asn	Gly	Asn	Leu	Glu	Lys	Thr	Thr	Tyr	Lys	Leu	Pro	Glu	Gly	Lys

780

775

```
Trp Asn Val Val Val Asn Ser Gln Lys Ala Gly Thr Glu Val Ile Glu 785 790 795 800
Thr Val Glu Gly Thr Ile Glu Leu Asp Pro Leu Ser Ala Tyr Val Leu
  805 810
                                                    815
Tyr Arg Glu Ser Glu Lys Asp Glu Leu
           820
<210> 35
<211> 460
<212> PRT
<213> Artificial Sequence
<220>
<223> synthetic
<400> 35
Met Arg Val Leu Leu Val Ala Leu Ala Leu Leu Ala Leu Ala Ser l\phantom{a} 5 \phantom{a} 10 \phantom{a} 15
Ala Thr Ser Ala Lys Tyr Leu Glu Leu Glu Glu Gly Gly Val Ile Met
        20
                            25
Gln Ala Phe Tyr Trp Asp Val Pro Ser Gly Gly Ile Trp Trp Asp Thr
     35
                          40
                                           45
lie Arg Gln Lys Ile Pro Glu Trp Tyr Asp Ala Gly Ile Ser Ala Ile
                   55
                                      60
Trp Ile Pro Pro Ala Ser Lys Gly Met Ser Gly Gly Tyr Ser Met Gly
                             75
                70
Tyr Asp Pro Tyr Asp Tyr Phe Asp Leu Gly Glu Tyr Tyr Gln Lys Gly 85 90
The Val Glu The Arg Phe Gly Ser Lys Gln Glu Leu Ile Asn Met Ile
        100
                            105
Asn Thr Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile
                         120
                                           125
 115
Asn His Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Gly Asp
 130 135 140
Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala 145 \, 150 \, 155 \, 160
Asn Tyr Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly
             165 170
                                                 175
Thr Phe Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln
          180
                         185 190
Tyr Trp Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser 195 200 205
Ile Gly Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala 210 215 220
Trp Val Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly 225 230 235 240
Glu Tyr Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Ser
            245
                              250 255
Ser Gly Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Ala
          260 265
                                          270
Ala Phe Asp Asn Lys Asn Ile Pro Ala Leu Val Glu Ala Leu Lys Asn 275 280 285
Gly Gly Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val
                           300
               295
Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala
                  310
                                 315
```

```
Phe Ile Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr
                                  330
               325
Glu Glu Trp Leu Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His
                             345
          340
Asp Asn Leu Ala Gly Gly Ser Thr Ser Ile Val Tyr Tyr Asp Ser Asp
                       360
                                            365
    355
Glu Met Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile
                    375
                                        380
Thr Tyr lie Asn Leu Gly Ser Ser Lys Val Gly Arg Trp Val Tyr Val
                  39Ô
                                    395
Pro Lys Phe Ala Gly Ala Cys Ile His Glu Tyr Thr Gly Asn Leu Gly
                           410
            405
Gly Trp Val Asp Lys Tyr Val Tyr Ser Ser Gly Trp Val Tyr Leu Glu
420 430
          420
Ala Pro Ala Tyr Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp 435 440 445
Ser Tyr Cys Gly Val Gly Ser Glu Lys Asp Glu Leu
<210> 36
<211> 718
<212> PRT
<213> Artificial Sequence
<223> synthetic
<400> 36
Met Arg Val Leu Leu Val Ala Leu Ala Leu Leu Ala Leu Ala Ala Ser
                                 10
Ala Thr Ser Met Glu Thr Ile Lys Ile Tyr Glu Asn Lys Gly Val Tyr
                             25
                                                30
Lys Val Val Ile Gly Glu Pro Phe Pro Pro Ile Glu Phe Pro Leu Glu
                         40
                                             45
Gln Lys Ile Ser Ser Asn Lys Ser Leu Ser Glu Leu Gly Leu Thr Ile
                     55
                                         60
 50
Val Gln Gln Gly Asn Lys Val Ile Val Glu Lys Ser Leu Asp Leu Lys
                 70
Glu His The The Gly Leu Gly Glu Lys Ala Phe Glu Leu Asp Arg Lys
85 90 95
                                 90
             8.5
Arg Lys Arg Tyr Val Met Tyr Asn Val Asp Ala Gly Ala Tyr Lys Lys 100 105 110
Tyr Gln Asp Pro Leu Tyr Val Ser Ile Pro Leu Phe Ile Ser Val Lys
                         120
                               125
Asp Gly Val Ala Thr Gly Tyr Phe Phe Asn Ser Ala Ser Lys Val Ile
                               140
                      135
Phe Asp Val Gly Leu Glu Glu Tyr Asp Lys Val Ile Val Thr Ile Pro
145 150 155 160
               150
Glu Asp Ser Val Glu Phe Tyr Val Ile Glu Gly Pro Arg Ile Glu Asp
165 170 175
            165
Val Leu Glu Lys Tyr Thr Glu Leu Thr Gly Lys Pro Phe Leu Pro Pro
                                         190
                   185
Met Trp Ala Phe Gly Tyr Met Ile Ser Arg Tyr Ser Tyr Tyr Pro Gln
                                          205
                          200
      195
Asp Lys Val Val Glu Leu Val Asp Ile Met Gln Lys Glu Gly Phe Arg
                   215
Val Ala Gly Val Phe Leu Asp Ile His Tyr Met Asp Ser Tyr Lys Leu
```

225					230					235			_		240
Phe	Thr	Trp	His	Pro 245	Tyr	Arg	Phe	Pro	Glu 250	Pro	Lys	Lys	Leu	Ile 255	Asp
Glu	Leu	His	Lys 260		Asn	Val	Lys	Leu 265		Thr	Ile	Val	Asp 270	His	Gly
Ile	Arg	Val 275		Gln	Asn	Tyr	Ser 280	Pro	Phe	Leu	Ser	Gly 285	Met	Gly	Lys
	290					295					300			Trp	
Gly 305	Thr	Thr	Val	Tyr	Pro 310	Asp	Phe	Phe	Arg	Glu 315	Asp	Thr	Arg	Glu	Trp 320
				325					330					Gly 335	
Trp	Leu	Asp	340					345					350	Glu	
Arg	Asp	355					360					365		Arg	
Val	Thr 370					375					380			Lys	
Val 385	•				390					395				Ala	400
Ala				405					410					Phe 415	
Leu			420					425					430	Ile	
Thr	-	435					440					445		Leu	
	450					455					460			Cys	
465					470					475				Ser	480
Asp				485					490					Phe 495	
Arg			500					505					510	Phe	
		515					520					525		Arg	
-	530					535					540			Glu	
545					550					555				Asp	560
-				565					570					Leu 575	
Tyr	Ala	Pro	580					585					590	Leu	
,	Gly	595	-	-		-	600					605		Gly	
	610		-			615					620				Gly
625					630					635				Glu	640
				645					650					Asn 655	
	•		660					665					670	Thr	
	-	675					680					685		Ile	
Val	Glu	Lys	Thr	Met	Gln	Asn	Thr	Tyr	Val	Ala	Lys	Пе	Asn	Gln	гуs

```
695
    690
                                            700
Ile Arg Gly Lys Ile Asn Leu Glu Ser Glu Lys Asp Glu Leu
                     710
<210> 37
<211> 1434
<212> DNA
<213> Thermotoga maritima
<400> 37
atgaaagaaa cogotgotgo taaattogaa ogocagoaca tggacagooc agatotgggt 60
accetogitge caegeogite catggoegag tietteeegg agateeegaa gateeagite 120
gagggcaagg agtocaccaa cocgetegee tteegettet aegaccegaa egaggtgate 180
gacqqcaagc cgctcaagga ccacctcaag ttctccgtgg ccttctggca caccttcgtg 240
aacgagegee gegaeeegtt eggegaeeeg aeegeegage geeegtggaa eegettetee 300
gaccogatgg acaaggcott ogocogogtg gacgcootot togagttotg ogagaagoto 360
aacatcgagt acttetgett eeacgacege gacategeee eggagggeaa gaceeteege 420
gagaccaaca agateetega caaggtggtg gagegeatea aggagegeat gaaggaetee 480
aacgtgaage teetetgggg caeegecaae etetteteee accegegeta catgeaegge 540
gaagacacca cotgotoogo ogacqtqtto gootacqcoq coqoocaqqt gaagaaqqco 600
ctagagatca ccaaggaget gggeggegag ggetaegtgt tetggggegg eegegaggge 660
tacgagacco tootcaacao ogacotoggo otggagotgg agaacotogo cogottooto 720
cgcatggccg tggagtacgc caagaagate ggcttcaccg gccagttcct categagccg 780
aageegaagg ageegaeeaa geaceagtae gaettegaeg tggeeaeege etaegeette 840
ctcaagaacc acggcctcga cgagtacttc aagttcaaca tcgaggccaa ccacgccacc 900
ctogooggoo acacettoca geacgagetg egeatggeee geatectegg caagetegge 960
tecategacy ceaaceaggy egaceteete eteggetggg acacegacea gttecegace 1020
aacatetacg acaccacet egecatgtae gaggtgatea aggeeggegg etteaceaag 1080
ggcggcctca acttcgacgc caaggtgcgc cgcgcctcct acaaggtgga ggacctcttc 1140
atoggecaca tegeoggeat ggacacette geeetegget teaagatege etacaagete 1200
gccaaggacg gcgtgttcga caagttcatc gaggagaagt accgctcctt caaggagggc 1260
atcagcaagg agatcgtgga gggcaagacc gacttcgaga agctggagga gtacatcatc 1320
gacaaggagg acatcgaget geegteegge aageaggagt acetggagte ceteeteaac 1380
rectacateg tgaagaceat egeegagetg egeteegaga aggacgaget gtga
<210> 38
<211> 477
<212> PRT
<213> Thermotoga maritima
<400> 38
Met Lys Glu Thr Ala Ala Ala Lys Phe Glu Arg Gln His Met Asp Ser
                                   10
Pro Asp Leu Gly Thr Leu Val Pro Arg Gly Ser Met Ala Glu Phe Phe
            20
                                25
Pro Glu Ile Pro Lys Ile Gln Phe Glu Gly Lys Glu Ser Thr Asn Pro
       35
                            40
                                                45
Leu Ala Phe Arg Phe Tyr Asp Pro Asn Glu Val Ile Asp Gly Lys Pro
                        55
Leu Lys Asp His Leu Lys Phe Ser Val Ala Phe Trp His Thr Phe Val
Asn Glu Gly Arg Asp Pro Phe Gly Asp Pro Thr Ala Glu Arg Pro Trp
Asn Arg Phe Ser Asp Pro Met Asp Lys Ala Phe Ala Arg Val Asp Ala
           100
                               105
                                                    110
Leu Phe Glu Phe Cys Glu Lys Leu Asn Ile Glu Tyr Phe Cys Phe His
```

```
Asp Arg Asp Ile Ala Pro Glu Gly Lys Thr Leu Arg Glu Thr Asn Lys
           135
  130
Ile Leu Asp Lys Val Val Glu Arg Ile Lys Glu Arg Met Lys Asp Ser
                                                         160
                  150
                                      155
145
Asn Val Lys Leu Leu Trp Gly Thr Ala Asn Leu Phe Ser His Pro Arg
                                 170
               165
Tyr Met His Gly Ala Ala Thr Thr Cys Ser Ala Asp Val Phe Ala Tyr
                                               190
                             185
           180
Ala Ala Ala Gln Val Lys Lys Ala Leu Glu Ile Thr Lys Glu Leu Gly
                                            205
                         200
       195
Gly Glu Gly Tyr Val Phe Trp Gly Gly Arg Glu Gly Tyr Glu Thr Leu
                    215
                                         220
Leu Asn Thr Asp Leu Gly Leu Glu Leu Glu Asn Leu Ala Arg Phe Leu
                   230
                                    235
Arg Met Ala Val Glu Tyr Ala Lys Lys Ile Gly Phe Thr Gly Gln Phe
                                                     255
               245
                                 250
Leu Ile Glu Pro Lys Pro Lys Glu Pro Thr Lys His Gln Tyr Asp Phe
                                                270
           260
                              265
Asp Val Ala Thr Ala Tyr Ala Phe Leu Lys Asn His Gly Leu Asp Glu
     275
                          280
                                           285
Tyr Phe Lys Phe Asn Ile Glu Ala Asn His Ala Thr Leu Ala Gly His
                      295
                                          300
Thr Phe Gln His Glu Leu Arg Met Ala Arg Ile Leu Gly Lys Leu Gly
                                      315
                 310
Ser Ile Asp Ala Asn Gln Gly Asp Leu Leu Ely Trp Asp Thr Asp
                                  330
                                                     335
               325
Gin Phe Pro Thr Asn Ile Tyr Asp Thr Thr Leu Ala Met Tyr Glu Val
                             345
                                                 350
          340
Ile Lys Ala Gly Gly Phe Thr Lys Gly Gly Leu Asn Phe Asp Ala Lys
                                              365
                          360
       355
Val Arg Arg Ala Ser Tyr Lys Val Glu Asp Leu Phe Ile Gly His Ile
                       375
   370
Ala Gly Met Asp Thr Phe Ala Leu Gly Phe Lys Ile Ala Tyr Lys Leu
                                      395
                   390
Ala Lys Asp Gly Val Phe Asp Lys Phe Ile Glu Glu Lys Tyr Arg Ser
                                410
               405
Phe Lys Glu Gly Ile Gly Lys Glu Ile Val Glu Gly Lys Thr Asp Phe
                                                430
                           425
          420
Glu Lys Leu Glu Glu Tyr Ile Ile Asp Lys Glu Asp Ile Glu Leu Pro
                                            445
       435
                          440
Ser Gly Lys Gln Glu Tyr Leu Glu Ser Leu Leu Asn Ser Tyr Ile Val
                                     460
                     455
Lys Thr Ile Ala Glu Leu Arg Ser Glu Lys Asp Glu Leu
                                     475
<210> 39
<211> 1434
<212> DNA
<213> Thermotoga neapolitana
```

```
atgaaagaaa ccgctgctgc taaattcgaa cgccagcaca tggacagccc agatctgggt 60
accetggtge caegoggtte catggeogag ttettecegg agatecegaa ggtgeagtte 120
gagggcaagg agtccaccaa cocgotogoo ttcaagttot acgaccogga ggagatcato 180
gacggcaage egeteaagga ceaceteaag tteteegtgg cettetggca cacettegtg 240
aacqagggcc gcgacccgtt cggcgacccg accgccgacc gcccgtggaa ccgctacacc 300
gaccogatgg acaaggcott cgcccgcgtg gacgccotct tcgagttctg cgagaagctc 360
```

<400> 39

```
aacategagt acticigett ecaegacege gaeategeee eggagggeaa gaeeeteege 420
gagaccaaca agatectega caaqqtqqtq qaqeqeatea aggageqeat gaaggaetee 480
aacqtqaago teetetgggg cacegecaae etetteteee accegegeta catgeaegge 540
googocacca congonocyc ogacquette goonacqueg cogoccaggi gaagaaggoo 600
ctggagatca ccaaggaget gggeggegag ggetaegtgt tetggggegg cegegaggge 660
tacqagacco tootcaacac egacetegge ttegagetgg agaacetege eegetteete 720
cgcatggccg tggactacgc caagcgcatc ggcttcaccg gccagttcct catcgagccg 780
aageegaagg ageegaeeaa geaeeagtae gaettegaeg tggeeaeege etaegeette 840
ctcaagtccc acggcctcga cgagtacttc aagttcaaca tcgaggccaa ccacgccacc 900
ctcgccggcc acacetteca geacgagetg egeatggeec geatectegg caagetegge 960
recategacy ecaaceaggy egaceteete eteggetggg acacegacea gtteecgace 1020
aacgigiacg acaccaccit cgccaigtac gaggigatca aggccggcgg ciicaccaag 1080
ggcggcctca acttcgacgc caaggtgcgc cgcgcctcct acaaggtgga ggacctcttc 1140
atoggocaca togooggoat ggacacetto goodtogget toaaggtgge ctacaagete 1200
greaaggacg gcgtgctcga caagttcatc gaggagaagt accgctcctt ccgcgagggc 1260
atoggoogog acatogtoga gggcaaggtg gacttogaga agotggagga gtacatoato 1320
gacaaggaga ccatcgaget geegteegge aagcaggagt acctggagte ceteateaac 1380
tectacateg tgaagaceat eetggagetg egeteegaga aggaegaget gtga
<210> 40
<211> 477
<212> PRT
<213> Thermotoga neapolitana
<400> 40
Met Lys Glu Thr Ala Ala Ala Lys Phe Glu Arg Gln His Met Asp Ser
                                   10
1
Pro Asp Leu Gly Thr Leu Val Pro Arg Gly Ser Met Ala Glu Phe Phe
            20
                                25
Pro Glu Ile Pro Lys Val Gln Phe Glu Gly Lys Glu Ser Thr Asn Pro
                                                45
        35
                            40
Leu Ala Phe Lys Phe Tyr Asp Pro Glu Glu Ile Ile Asp Gly Lys Pro
                                            60
                        55
    50
Leu Lys Asp His Leu Lys Phe Ser Val Ala Phe Trp His Thr Phe Val
                                        75
                    70
Asn Glu Gly Arg Asp Pro Phe Gly Asp Pro Thr Ala Asp Arg Pro Trp
                                    90
                85
Asn Arg Tyr Thr Asp Pro Met Asp Lys Ala Phe Ala Arg Val Asp Ala
                                                    110
                               105
            100
Leu Phe Glu Phe Cys Glu Lys Leu Asn Ile Glu Tyr Phe Cys Phe His
                                                125
                            120
       115
Asp Arg Asp Ile Ala Pro Glu Gly Lys Thr Leu Arg Glu Thr Asn Lys
                                            140
                        1.35
Ile Leu Asp Lys Val Val Glu Arg Ile Lys Glu Arg Met Lys Asp Ser
                                                            160
                                        155
                    150
145
Asn Val Lys Leu Leu Trp Gly Thr Ala Asn Leu Phe Ser His Pro Arg
                                   170
                165
Tyr Met His Gly Ala Ala Thr Thr Cys Ser Ala Asp Val Phe Ala Tyr
                                                    190
           180
                               185
Ala Ala Gln Val Lys Lys Ala Leu Glu Ile Thr Lys Glu Leu Gly
                                                205
        195
                            200
Gly Glu Gly Tyr Val Phe Trp Gly Gly Arg Glu Gly Tyr Glu Thr Leu
                                            220
                        215
    210
Leu Asn Thr Asp Leu Gly Phe Glu Leu Glu Asn Leu Ala Arg Phe Leu
                                        235
                    230
Arg Met Ala Val Asp Tyr Ala Lys Arg Ile Gly Phe Thr Gly Gln Phe
                                    250
                245
Leu Ile Glu Pro Lys Pro Lys Glu Pro Thr Lys His Gln Tyr Asp Phe
```

```
265
           260
Asp Val Ala Thr Ala Tyr Ala Phe Leu Lys Ser His Gly Leu Asp Glu
                           280
                                               285
       275
Tyr Phe Lys Phe Asn Ile Glu Ala Asn His Ala Thr Leu Ala Gly His
                                            300
                       295
   290
Thr Phe Gln His Glu Leu Arg Met Ala Arg Ile Leu Gly Lys Leu Gly
                                        315
305
                    310
Ser Ile Asp Ala Asn Gln Gly Asp Leu Leu Gly Trp Asp Thr Asp
                                                        335
                                    330
               325
Gin Phe Pro Thr Asn Val Tyr Asp Thr Thr Leu Ala Met Tyr Glu Val
                                345
           340
lle Lys Ala Gly Gly Phe Thr Lys Gly Gly Leu Asn Phe Asp Ala Lys
                            360
                                                365
       355
Val Arg Arg Ala Ser Tyr Lys Val Glu Asp Leu Phe Ile Gly His Ile
                        375
   370
Ala Gly Met Asp Thr Phe Ala Leu Gly Phe Lys Val Ala Tyr Lys Leu
                                        395
                    390
Val Lys Asp Gly Val Leu Asp Lys Phe Ile Glu Glu Lys Tyr Arg Ser
                                    410
                405
Phe Arg Glu Gly Ile Gly Arg Asp Ile Val Glu Gly Lys Val Asp Phe
                                425
           420
Glu Lys Leu Glu Glu Tyr Ile Ile Asp Lys Glu Thr Ile Glu Leu Pro
                                                445
                            440
Ser Gly Lys Gln Glu Tyr Leu Glu Ser Leu Ile Asn Ser Tyr Ile Val
                                           460
                        455
   450
Lys Thr Ile Leu Glu Leu Arg Ser Glu Lys Asp Glu Leu
                    470
<210> 41
<211> 1435
<212> DNA
<213> Thermotoga maritima
<400> 41
atgggcaçca gccatcatca tcatcatcac agcagcggcc tggtgccgcg cggcagccat 60
atggctagca tgactggtgg acagcaaatg ggtcggatcc ccatggccga gttcttcccg 120
gagatecega agatecagtt egagggeaag gagtecacea accegetege etteegette 180
tacgaccega acgaggigat egaeggeaag eegeteaagg accaceteaa giteteegig 240
geettetgge acacettegt gaacgagge egegaceegt teggegacee gacegeegag 300
egocogtgga acogettete egacoegatg gacaaggeet tegeoogegt ggacgeeete 360
ttcgagttct gcgagaaget caacatcgag tacttetget tecaegaceg cgacatecee 420
cggagggcaa gacceteege gagaceaaca agateetega caaggtggtg gagegeatea 480
aggagegeat gaaggaetee aacgtgaage teetetgggg caeegecaae etetteteee 540
accogogota catgoacogo geogocacca cotoctocogo egacototto geotacogog 600
ccgcccaggt gaagaaggcc ctggagatca ccaaggagct gggcggcgag ggctacgtgt 660
tetggggegg cegegagge tacgagacce tectcaacac egacetegge etggagetgg 720
agaacetege edgetteete egeatggeeg togagtaege caagaagate ggetteaceg 780
gocagitect categageeg aageegaagg ageegaceaa geaceagiae gettegaegi 840
ggccaccgcc tacgcettec teaagaacca eggcetegae gagtaettea agttcaacat 900
egaggeeaac cacqccacce tegeoggeea cacettecag cacgagetge geatggeecg 960
catectegge aagetegget ceategacge caaccaggge gacetectee teggetggga 1020
caccgaccag ttcccgacca acatctacga caccaccctc gccatgtacg aggtgatcaa 1080
ggccggcgc ttcaccaagg gcggcctcaa cttcgacgcc aaggtgcgcc gcgcctccta 1140
caaggiggag gaccicitca toggocacat ogcoggoatg gacaccitog coologgoit 1200
caagatogoo tacaagotog ccaaggaogg ogtgttogac aagttoatog aggagaagta 1260
ccgctccttc aaggagggca tcggcaagga gatcgtggag ggcaagaccg acttcgagaa 1320
gctggaggag tacatcatcg acaaggagga catcgagctg ccgtccggca agcaggagta 1380
```

cctggagtcc ctcctcaact cctacatcgt gaagaccatc gccgagctgc gctga <210> 42 <211> 473 <212> PRT <213> Thermotoga maritima <400> 42 Met Gly Ser Ser His His His His His Ser Ser Gly Leu Val Pro Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg Ile Pro Met Ala Glu Phe Phe Pro Glu Ile Pro Lys Ile Gln Phe Glu Gly Lys Glu Ser Thr Asn Pro Leu Ala Phe Arg Phe Tyr Asp Pro Asn Glu Val Ile Asp Gly Lys Pro Leu Lys Asp His Leu Lys Phe Ser Val Ala Phe Trp His Thr Phe Val Asn Glu Gly Arg Asp Pro Phe Gly Asp Pro Thr Ala Glu Arg Pro Trp Asn Arg Phe Ser Asp Pro Met Asp Lys Ala Phe Ala Arg Val Asp Ala Leu Phe Glu Phe Cys Glu Lys Leu Asn 115 120 125 Ile Glu Tyr Phe Cys Phe His Asp Arg Asp Ile Ala Pro Glu Gly Lys Thr Leu Arg Glu Thr Asn Lys Ile Leu Asp Lys Val Val Glu Arg Ile Lys Glu Arg Met Lys Asp Ser Asn Val Lys Leu Leu Trp Gly Thr Ala Asn Leu Phe Ser His Pro Arg Tyr Met His Gly Ala Ala Thr Thr Cys Ser Ala Asp Val Phe Ala Tyr Ala Ala Ala Gln Val Lys Lys Ala Leu Glu Ile Thr Lys Glu Leu Gly Gly Glu Gly Tyr Val Phe Trp Gly Gly Arg Glu Gly Tyr Glu Thr Leu Leu Asn Thr Asp Leu Gly Leu Glu Leu Glu Asn Leu Ala Arg Phe Leu Arg Met Ala Val Glu Tyr Ala Lys Lys Ile Gly Phe Thr Gly Gln Phe Leu Ile Glu Pro Lys Pro Lys Glu Pro Thr Lys His Gln Tyr Asp Phe Asp Val Ala Thr Ala Tyr Ala Phe Leu Lys Asn His Gly Leu Asp Glu Tyr Phe Lys Phe Asn Ile Glu Ala Asn His Ala Thr Leu Ala Gly His Thr Phe Gln His Glu Leu Arg Met Ala Arg Ile Leu Gly Lys Leu Gly Ser Ile Asp Ala Asn Gln Gly Asp Leu 325 330 335 Leu Leu Gly Trp Asp Thr Asp Gln Phe Pro Thr Asn Ile Tyr Asp Thr Thr Leu Ala Met Tyr Glu Val Ile Lys Ala Gly Gly Phe Thr Lys Gly Gly Leu Asn Phe Asp Ala Lys Val Arg Arg Ala Ser Tyr Lys Val Glu Asp Leu Phe Ile Gly His Ile Ala Gly Met Asp Thr Phe Ala Leu Gly

Page 44

```
Phe Lys Ile Ala Tyr Lys Leu Ala Lys Asp Gly Val Phe Asp Lys Phe
                405
                                   410
Ile Glu Glu Lys Tyr Arg Ser Phe Lys Glu Gly Ile Gly Lys Glu Ile
                                                    430
                                425
Val Glu Gly Lys Thr Asp Phe Glu Lys Leu Glu Glu Tyr Ile Ile Asp
                            440
                                                445
Lys Glu Asp Ile Glu Leu Pro Ser Gly Lys Gln Glu Tyr Leu Glu Ser
                       455
                                            460
Leu Leu Asn Ser Tyr Ile Val Lys Thr Ile Ala Glu Leu Arg
                    470
                                        475
<210> 43
<211> 1436
<212> DNA
<213> Thermotoga neapolitana
<400> 43
atgggcagca gccatcatca tcatcatcac agcagcggcc tggtgccgcg cggcagccat 60
atggctagca tgactggtgg acagcaaatg ggtcggatcc ccatggccga gttcttcccg 120
gagatecega aggtgeagtt egagggeaag gagteeacea accegetege etteaagtte 180
tacqaccogg aggagateat egacggeaag cegeteaagg accaecteaa gtteteegtg 240
gccttctggc acaccttcgt gaacgagggc cgcgacccgt tcggcgaccc gaccgccgac 300
cgcccgtgga accgctacac cgacccgatg gacaaggeet tegecegegt ggacgccete 360
ttcgagttct gcgagaagct caacatcgag tacttctgct tccacgaccg cgacatcccc 420
cggagggcaa gacceteege gagaceaaca agateetega caaggtggtg gagegeatea 480
aggagegeat gaaggaetee aaegtgaage teetetgggg cacegeeaae etetteteee 540
accogogota catgoacggo googocacca cotgotocgo cgacqtqtto gootacgcog 600
ccgcccaggt gaagaaggcc ctggagatca ccaaggagct gggcggcgag ggctacgtgt 660
totggggcgg cogogagggo tacgagacco tootcaacac ogacotoggo ttogagotgg 720
agaacetege eegetteete egeatggeeg tggactaege caagegeate ggetteaceg 780
gccagttcct catcgagccg aagccgaagg agccgaccaa gcaccagtac gacttcgacg 840
tggccaccgc ctacgccttc ctcaagtccc acggcctcga cgagtacttc aagttcaaca 900
togaggodaa coacgodaco otogooggod acacettoca goacgagetg ogdatggood 960
gcatcetegg caagetegge tecategacg ceaaceaggg egaceteete eteggetggg 1020
acaccgacca gttcccgacc aacgtgtacg acaccacct cgccatgtac gaggtgatca 1080
aggcoggogg cttcaccaag ggoggoctca acttcgacgc caaggtgcgc cgcgcctcct 1140
acaaggtgga ggacetette ateggecaca tegeoggeat ggacacette geeetegget 1200
tcaaggtggc ctacaagctc gtgaaggacy gcgtgctcga caagttcatc gaggagaagt 1260
accyctectt ecgegagge ateggeegeg acategtgga gggcaaggtg gaettegaga 1320
agctggagga gtacatcatc gacaaggaga ccatcgagct gccgtccggc aagcaggagt 1380
acctggagtc cctcatcaac tcctacatcg tgaagaccat cctggagctg cgctga
<210> 44
<211> 478
<212> PRT
<213> Thermotoga neapolitana
<400> 44
Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro
                                    10
                 5
Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg
                                25
           20
Ile Pro Met Ala Glu Phe Phe Pro Glu Ile Pro Lys Val Gln Phe Glu
                                                45
                            40
Gly Lys Glu Ser Thr Asn Pro Leu Ala Phe Lys Phe Tyr Asp Pro Glu
                                            60
                       55
Glu Ile Ile Asp Gly Lys Pro Leu Lys Asp His Leu Lys Phe Ser Val
```

```
70
Ala Phe Trp His Thr Phe Val Asn Glu Gly Arg Asp Pro Phe Gly Asp
                                   90
             8.5
Pro Thr Ala Asp Arg Pro Trp Asn Arg Tyr Thr Asp Pro Met Asp Lys
100 105 110
Ala Phe Ala Arg Val Asp Ala Leu Phe Glu Phe Cys Glu Lys Leu Asn 115 120 125
Ile Glu Tyr Phe Cys Phe His Asp Arg Asp Ile Ala Pro Glu Gly Lys
130 135 140
Thr Leu Arg Glu Thr Asn Lys Ile Leu Asp Lys Val Val Glu Arg Ile 145 150 155 160
Lys Glu Arg Met Lys Asp Ser Asn Val Lys Leu Leu Trp Gly Thr Ala 165 170 175
Asn Leu Phe Ser His Pro Arg Tyr Met His Gly Ala Ala Thr Thr Cys
180 185 190
         180
Ser Ala Asp Val Phe Ala Tyr Ala Ala Ala Gln Val Lys Lys Ala Leu
195 200 205
Glu Ile Thr Lys Glu Leu Gly Gly Glu Gly Tyr Val Phe Trp Gly Gly
                                           220
                      215
Arg Glu Gly Tyr Glu Thr Leu Leu Asn Thr Asp Leu Gly Phe Glu Leu 225 230 240
225
Glu Asn Leu Ala Arg Phe Leu Arg Met Ala Val Asp Tyr Ala Lys Arg
245 250 255
Ile Gly Phe Thr Gly Gln Phe Leu Ile Glu Pro Lys Pro Lys Glu Pro 260 270
Thr Lys His Gln Tyr Asp Phe Asp Val Ala Thr Ala Tyr Ala Phe Leu 275 280 285
Lys Ser His Gly Leu Asp Glu Tyr Phe Lys Phe Asn Ile Glu Ala Asn
 290
                                         300
                       295
His Ala Thr Leu Ala Gly His Thr Phe Gln His Glu Leu Arg Met Ala 305 310 315 320
Arg Ile Leu Gly Lys Leu Gly Ser Ile Asp Ala Asn Gln Gly Asp Leu 325 330 335
Leu Leu Gly Trp Asp Thr Asp Gln Phe Pro Thr Asn Val Tyr Asp Thr 340 345 350
Thr Leu Ala Met Tyr Glu Val Ile Lys Ala Gly Gly Phe Thr Lys Gly
                                                  365
 355 360
Gly Leu Asn Phe Asp Ala Lys Val Arg Arg Ala Ser Tyr Lys Val Glu
370 380
                    375
Asp Leu Phe Ile Gly His Ile Ala Gly Met Asp Thr Phe Ala Leu Gly 385 390 395 400
Phe Lys Val Ala Tyr Lys Leu Val Lys Asp Gly Val Leu Asp Lys Phe 405 410 415
Ile Glu Glu Lys Tyr Arg Ser Phe Arg Glu Gly Ile Gly Arg Asp Ile 420 430
Val Glu Gly Lys Val Asp Phe Glu Lys Leu Glu Glu Tyr Ile Ile Asp
435 440 445
                        440
 435
Lys Glu Thr Ile Glu Leu Pro Ser Gly Lys Gln Glu Tyr Leu Glu Ser
 450 455
Leu Ile Asn Ser Tyr Ile Val Lys Thr Ile Leu Glu Leu Arg
465 470
```

<210> 45

<211> 1095

<212> PRT

<213> Aspergillus shirousami

<400> 45 Ala Thr Pro Ala Asp Trp Arg Ser Gln Ser Ile Tyr Phe Leu Leu Thr $1 \hspace{1cm} 1 \hspace{1cm} 15$ Asp Arg Phe Ala Arg Thr Asp Gly Ser Thr Thr Ala Thr Cys Asn Thr 25 30 Ala Asp Gln Lys Tyr Cys Gly Gly Thr Trp Gln Gly Ile Ile Asp Lys 40 Leu Asp Tyr Ile Gln Gly Met Gly Phe Thr Ala Ile Trp Ile Thr Pro 55 50 Val Thr Ala Gln Leu Pro Gln Thr Thr Ala Tyr Gly Asp Ala Tyr His 65 70 75 80 70 Gly Tyr Trp Gln Gln Asp Ile Tyr Ser Leu Asn Glu Asn Tyr Gly Thr 85 90 95 8.5 Ala Asp Asp Leu Lys Ala Leu Ser Ser Ala Leu His Glu Arg Gly Met 105 110 100 Tyr Leu Met Val Asp Val Val Ala Asn His Met Gly Tyr Asp Gly Ala 115 125 Gly Ser Ser Val Asp Tyr Ser Val Phe Lys Pro Phe Ser Ser Gln Asp 135 Tyr Phe His Pro Phe Cys Phe Ile Gln Asn Tyr Glu Asp Gln Thr Gln 145 150 155 160 Val Glu Asp Cys Trp Leu Gly Asp Asn Thr Val Ser Leu Pro Asp Leu 165 170 175 Asp Thr Thr Lys Asp Val Val Lys Asn Glu Trp Tyr Asp Trp Val Gly 185 190 Ser Leu Val Ser Asn Tyr Ser Ile Asp Gly Leu Arg Ile Asp Thr Val 200 205 195 Lys His Val Gln Lys Asp Phe Trp Pro Gly Tyr Asn Lys Ala Ala Gly 215 220 Val Tyr Cys Ile Gly Glu Val Leu Asp Val Asp Pro Ala Tyr Thr Cys 230 235 240 Pro Tyr Gln Asn Val Met Asp Gly Val Leu Asn Tyr Pro Ile Tyr Tyr 245 250 Pro Leu Leu Asn Ala Phe Lys Ser Thr Ser Gly Ser Met Asp Asp Leu 260 265 Tyr Asn Met Ile Asn Thr Val Lys Ser Asp Cys Pro Asp Ser Thr Leu 275 280 285 Leu Gly Thr Phe Val Glu Asn His Asp Asn Pro Arg Phe Ala Ser Tyr 290 295 300 Thr Asn Asp Ile Ala Leu Ala Lys Asn Val Ala Ala Phe Ile Ile Leu 305 310 315 310 315 Asn Asp Gly Ile Pro Ile Ile Tyr Ala Gly Gln Glu Gln His Tyr Ala 330 335 Gly Gly Asn Asp Pro Ala Asn Arg Glu Ala Thr Trp Leu Ser Gly Tyr 345 350 340 Pro Thr Asp Ser Glu Leu Tyr Lys Leu Ile Ala Ser Ala Asn Ala Ile 360 365 355 Arg Asn Tyr Ala Ile Ser Lys Asp Thr Gly Phe Val Thr Tyr Lys Asn 370 380 Trp Pro Ile Tyr Lys Asp Asp Thr Thr Ile Ala Met Arg Lys Gly Thr 390 395 Asp Gly Ser Gln Ile Val Thr Ile Leu Ser Asn Lys Gly Ala Ser Gly 405 410 Asp Ser Tyr Thr Leu Ser Leu Ser Gly Ala Gly Tyr Thr Ala Gly Gln 420 425 430 425 420 Gln Leu Thr Glu Val Ile Gly Cys Thr Thr Val Thr Val Gly Ser Asp 440 Gly Asn Val Pro Val Pro Met Ala Gly Gly Leu Pro Arg Val Leu Tyr

	450					455					460				
Pro 465		Glu	Lys	Leu	Ala 470	Gly	Ser	Lys	Ile	Cys 475	Ser	Ser	Ser	Lys	Pro 480
	Thr	Leu	Asp	Ser 485	Trp	Leu	Ser	Asn	Glu 490	Ala	Thr	Val	Ala	Arg 495	Thr
Ala	Ile	Leu	Asn 500		Ile	Gly	Ala	Asp 505	Gly	Ala	Trp	Val	Ser 510	Gly	Ala
Asp	Ser	Gly 515	Ile	Va l	Val	Ala	Ser 520	Pro	Ser	Thr	Asp	Asn 525	Pro	Asp	Tyr
	530		-		Arg	535					540				
545					Gly 550					555					560
				565	Ala				570					575	
•			580		Gly			585					590		
		595			Ser		600					605			
	610				Met	615					620				
625					Ala 630					635					640
-				645	Ala				650					655	
			660		Gly			665					670		
-		675			Gly		680					685			
-	690				Ser	695					700				
705					Ser 710					715					720
	-			725	Asn Asp				730					735	
			740		Lys			745					750		
		755			Lys		760					765			
	770				Tyr	775					780				
785					790 Gln					795					800
				805	Ile				810					815	
	•		820		Ala			825					830		
		835			Ser		840					845			
•	850				His	855					860				
865					870 Gly					875					880
	-	-		885	Leu				890					895	
			900		Glu			905					910		
r:0	F 1 0	261	rrb	GTA	GIU		Jei	1110	J C 1				 J		-1-

```
920
                                                 925
Ala Ala Thr Ser Ala Ser Gly Thr Tyr Ser Ser Val Thr Val Thr Ser
 930
                        935
                                             940
Trp Pro Ser Ile Val Ala Thr Gly Gly Thr Thr Thr Thr Ala Thr Thr
                    950
                                         955
945
Thr Gly Ser Gly Gly Val Thr Ser Thr Ser Lys Thr Thr Thr Ala
                                     970
                965
Ser Lys Thr Ser Thr Thr Thr Ser Ser Thr Ser Cys Thr Thr Pro Thr
                                 985
                                                     990
            980
Ala Val Ala Val Thr Phe Asp Leu Thr Ala Thr Thr Tyr Gly Glu
                           1000
                                               1005
        995
Asn Ile Tyr Leu Val Gly Ser Ile Ser Gln Leu Gly Asp Trp Glu Thr
                        1015
                                            1020
   1010
Ser Asp Gly Ile Ala Leu Ser Ala Asp Lys Tyr Thr Ser Ser Asn Pro
                    1030
                                         1035
1025
Pro Trp Tyr Val Thr Val Thr Leu Pro Ala Gly Glu Ser Phe Glu Tyr
                                                         1055
                                    1050·
                1045
Lys Phe Ile Arg Val Glu Ser Asp Asp Ser Val Glu Trp Glu Ser Asp
                                                    1070
                                1065
            1060
Pro Asn Arg Glu Tyr Thr Val Pro Gln Ala Cys Gly Glu Ser Thr Ala
                                                 1085
                            1080
       1075
Thr Val Thr Asp Thr Trp Arg
    1090
                        1095
<210> 46
<211> 3285
<212> DNA
<213> Aspergillus shirousami
<400> 46
gccaccccgg ccgactggcg ctcccagtcc atctacttcc tcctcaccga ccgcttcgcc 60
ogcaccgacg gotocaccac egocacctgo aacacegoog accagaagta ctgcggcggc 120
acctograge gratratega caagetegae tacateragg gratgggett caregorate 180
tggatcacco eggtgacego ecageteceg cagaccaceg cetacggega egectaceae 240
ggctactggc agcaggacat ctactccctc aacgagaact acggcaccgc cgacgacctc 300
aaggoodtot ootoogoodt ooacgagege ggeatgtace teatggtgga egtggtggee 360
aaccacatgg gctacgacgg cgccggctcc tccgtggact actccgtgtt caagccgttc 420
tecteccagg actaetteca ecegitetge ticatecaga actaegagga ecagaeceag 480
grggaggact gerggeregg egacaacace grgreetee eggacerega caccaccaag 540
gacgtggtga agaacgagtg gtacgactgg gtgggctccc tcgtgtccaa ctactccatc 600 gacggcctcc gcatcgacac cgtgaagcac gtgcagaagg acttctggcc gggctacaac 660
aaggoogoog gogtgtactg categgogag gtgetegaeg tggaccegge ctacacetge 720
ccgtaccaga acgtgatgga cggcgtgctc aactacccga tctactaccc gctcctcaac 780
gccttcaagt ccacctccgg ctcgatggac gacctctaca acatgatcaa caccgtgaag 840
tecgactique eggactecae estectegge acettegtigg agaaceaega caaceegege 900
trogeorect acaccaacga categooote gecaagaacg tggccgcott cateatoote 960
aacgacggca teeegateat etaegeegge caggageage actaegeegg eggeaacgae 1020
ccggccaacc gcgaggccac ctggctctcc ggctacccga ccgactccga gctgtacaag 1080
ctcatcgcct ccgccaacgc catccgcaac tacgccatct ccaaggacac cggcttcgtg 1140
acctacaaga actggccgat ctacaaggac gacaccacca tcgccatgcg caagggcacc 1200
gacggetece agategtgae cateetetee aacaagggeg ceteeggega etectacace 1260
ctctccctct ccggcgccgg ctacaccgcc ggccagcagc tcaccgaggt gatcggctgc 1320
accaccqtqa ccqtqggctc cgacqgcaac qtqccqqtqc cgatqqccqq cgqcctcccq 1380
egegtgetet accegacega gaagetegee ggetecaaga tatgeteete etecaageeg 1440
gecacceteg acteetgget etecaacgag gecaccgtgg eeggcaccge cateetcaac 1500
aacateggeg eegaeggege etgggtgtee ggegeegaet eeggeategt ggtggeetee 1560
cogrecaceg acaaceegga ctacttetac acetggacee gegacteegg categtgete 1620
```

```
aagaccotog tggacctott cogcaacggo gacaccgaco tootetocac catogagoac 1680
tacateteet eccaggeeat cateeaggge gtgtecaace egteeggega ecteteetee 1740
ggcggcctcg gcgagccgaa gttcaacgtg gacgagaccg cctacgccgg ctcctggggc 1800
cgcccgcage gegacggeee ggeeeteege gecacegeea tgateggett eggeeagtgg 1860
ctectegaca acggetacae etecgeegee accgagateg tgtggceget egtgegeaac 1920
gacetetect acottogecea gractggaae cagacegget aegacetetg ggaggaggtg 1980
aacggeteet cettetteac categoogtg cageacegeg ceetegtgga gggeteegee 2040
ttegecaceg degtgggete etectgetee tggtgegaet cecaggeece geagateete 2100
toctacetee agreeteetg gaceggetee tacateeteg ccaacttega etectocege 2160
teeggeaagg acaceaacae ecteetegge teeateeaca cettegacce ggaggeegge 2220
tgcgacqact ccaccttcca gccgtgctcc ccgcgcgccc tcgccaacca caaggaggtg 2280
gragactect reegeteear oracaceete aacgaeggee teteegaete egaggeegtg 2340
geogragges getaceegga ggacteetas tacaaeggsa acceptggtt estetgsace 2400
ctogoogoog cogagoagot ctacgaogoo ototaccagt gggacaagoa gggotocotg 2460
gagateaceg aegitgiedet egacitetie aaggeeetet aeteeggege egeeacegge 2520
acctactect ectectecte cacetactee tecategtgt ecgeogtgaa gacettegee 2580
gacggetteg tgtecategt ggagaceeac geogeeteea aeggeteeet eteegageag 2640
troqueur cogaeggega egagetgree geoegegaee reacetggre etaegeegee 2700
ercercaceg ecaacaaceg eegcaactee grggrgeege egteerggg egagacetee 2760
geotectecg tgeegggeac etgegeegec aceteegeet eeggeaceta etecteegtg 2820
accettacct cettacette categtages accaseas caccase egecaccase 2880
accogetec geogeotiae etecacetec aagaceacea ceacegeete caagacetec 2940
accaccacct ectecacete etgeaceace ecgacegecg tggeegtgae ettegacete 3000
acceccacca coacctaceg egagaacate tacctegteg getecatete ecagetegge 3060
gactoggaga cotocgacgg categocote teegeegaca agtacacete etecaaceeg 3120
cogtggtacg tgaccgtgac cctcccggcc ggcgagtcct tcgagtacaa gttcatccgc 3180
atgaagtccq acqactccqt ggagtgggag tccgacccga accgcgagta caccgtgccg 3240
                                                                   3285
caggeetgeg gegagteeac egecacegtg accgaeacet ggege
<210> 47
<211> 679
<212> PRT
<213> Thermoanaerobacterium thermosaccharolyticum
<400> 47
Val Leu Ser Gly Cys Ser Asn Asn Val Ser Ser Ile Lys Ile Asp Arg
                                    10
                                                        15
Phe Asn Asn Ile Ser Ala Val Asn Gly Pro Gly Glu Glu Asp Thr Trp
                                25
            20
Ala Ser Ala Gln Lys Gln Gly Val Gly Thr Ala Asn Asn Tyr Val Ser
                            4.0
                                                45
Arg Val Trp Phe Thr Leu Ala Asn Gly Ala Ile Ser Glu Val Tyr Tyr
    50
                        55
                                            60
Pro Thr Ile Asp Thr Ala Asp Val Lys Glu Ile Lys Phe Ile Val Thr
                    70
                                        75
Asp Gly Lys Ser Phe Val Ser Asp Glu Thr Lys Asp Ala Ile Ser Lys
                                    90
                                                        95
                85
Val Glu Lys Phe Thr Asp Lys Ser Leu Gly Tyr Lys Leu Val Asn Thr
                                105
                                                    110
           100
Asp Lys Lys Gly Arg Tyr Arg Ile Thr Lys Glu Ile Phe Thr Asp Val
                            120
       115
Lys Arg Asn Ser Leu Ile Met Lys Ala Lys Phe Glu Ala Leu Glu Gly
                      135
                                          140
  130
Ser Ile His Asp Tyr Lys Leu Tyr Leu Ala Tyr Asp Pro His Ile Lys
                   150
                                       155
                                                            160
145
Asn Gln Gly Ser Tyr Asn Glu Gly Tyr Val Ile Lys Ala Asn Asn Asn
                                   170
                                                        175
               165
```

Glu Met Leu Met Ala Lys Arg Asp Asn Val Tyr Thr Ala Leu Ser Ser

													100		
Asn	Ile	Gly	180 Trp	Lys	Gly	Tyr	Ser	185 Ile	Gly	Tyr	Tyr	Lys	190 Val	Asn	Asp
		195					200					205			
116	210	inr	Asp	Leu	Asp	215	ASN	Lys	GIN	met	220	rys	uis	ıyı	АЅР
Ser 225	Ala	Arg	Gly	Asn	11e 230	Ile	Glu	Gly	Ala	Glu 235	Ile	Asp	Leu	Thr	Lys 240
	Ser	Glu	Phe	Glu 245	Ile	Val	Leu	Ser	Phe 250		Gly	Ser	Asp	Ser 255	
Ala	Ala	Lys	Thr 260		Leu	Glu	Thr	Leu 265		Glu	Asp	Tyr	Asn 270		Leu
Lys	Asn	Asn 275		Ile	Asp	Glu	Trp 280	Thr	Lys	Tyr	Суѕ	Asn 285	Thr	Leu	Asn
Asn	Phe 290		Gly	Lys	Ala	Asn 295	Ser	Leu	Tyr	Tyr	Asn 300	Ser	Met	Met	Ile
Leu 305	Lys	Ala	Ser	Glu	Asp 310	Lys	Thr	Asn	Lys	Gly 315	Ala	Tyr	Ile	Ala	Ser 320
Leu	Ser	Ile	Pro	Trp 325	Gly	Asp	Gly	Gln	Arg 330		Asp	Asn	Thr	Gly 335	
Tyr	Нis	Leu	Val 340	Trp	Ser	Arg	Asp	Leu 345	Tyr	His	Val	Ala	Asn 350	Ala	Phe
Ile	Ala	Ala 355		Asp	Val	Asp	Ser 360		Asn	Arg	Ser	Leu 365	Asp	Tyr	Leu
	370	Val			Asp	375					380				
385					Trp 390					395					400
				405	Tyr				410					415	
			420		Ala			425					430		
•		435			Glu		440					445			
	450				Gly	455					460				
465	-	-	-		Ser 470					475					480
-				485	Asp				490					495	
-		_	500		Tyr			505					510		
	-	515			Asn		520					525			
•	530				Pro	535					540				
Lys 545	Ser	Ala	Asp	Asp	Pro 550	Lys	Ile	Leu	Asn	Thr 555	Leu	Lys	Val	Val	Asp 560
Ser	Thr	Ile	Lys	Val 565	Asp	Thr	Pro	Lys	Gly 570	Pro	Ser	Trp	Tyr	Arg 575	Tyr
Asn	Hıs	Asp	Gly 580	Tyr	Gly	Glu	Pro	Ser 585	Lys	Thr	Glu	Leu	Tyr 590	His	Gly
	-	595			Leu		600					605			
-	610				Gly	615					620				
Glu 625	Lys	Phe	Ala	Asn	Glu 630	Gly	Gly	Ile	Ile	Ser 635	Glu	Gln	Val	Trp	Glu 640
Asp	Thr	Gly	Leu	Pro	Thr	Asp	Ser	Ala	Ser		Leu	Asn	Trp	Ala	His

```
645
                                    650
Ala Glu Tyr Val Ile Leu Phe Ala Ser Asn Ile Glu His Lys Val Leu
                                665
                                                    670
          660
Asp Met Pro Asp Ile Val Tyr
        675
<210> 48
<211> 2037
<212> DNA
<213> Thermoanaerobacterium thermosaccharolyticum
<2205
<223> synthetic
<400> 48
attactored actactoraa caacatatee tocatcaaga togacogett caacaacate 60
tecgeogtga acggeoeggg egaggaggae acetgggeet eegeceagaa geagggegtg 120
ggcaccgcca acaactacgt gtcccgcgtg tggttcaccc tcgccaacgg cgccatctcc 180
gaggigiact accegaceat egacacegee gacgigaagg agaicaagii categigace 240
gacqqcaagt cetteqtqte eqacqaqaee aaggacqcea tetecaaggt ggagaagtte 300
accgacaagt coctoggota caagetogtg aacaccgaca agaagggoog ctaccgcate 360
accaaggaaa tottcaccga cgtgaagcgc aactocotca toatgaaggc caagttcgag 420
gccctcgagg gctccatcca cgactacaag ctctacctcg cctacgaccc gcacatcaag 480
aaccaggget ectacaacga gggetacgtg atcaaggeca acaacaacga gatgetcatg 540
gccaagegeg acaaegtgta caeegeeete teetecaaca teggetggaa gggetaetee 600
ateggetact acaaggigaa egacateatg acegaceteg acgagaacaa geagatgace 660
aagcactacg actoogooog oggoaacato atogagggog oogagatoga ootoaccaag 720
aactocgagt togagatogt geteteette ggeggeteeg actocgagge egecaagace 780
occetegaga coeteggega ggactacaac aaceteaaga acaactacat egacgagtgg 840
accaagtact gcaacaccet caacaactte aacggcaagg ccaactecet etactacaac 900
tocatgatga tootcaaggo otocgaggac aagaccaaca agggcgcota categootec 960
ctotocatoc cgtggggcga cggccagcgc gacgacaaca ccggcggcta ccacctcgtg 1020
tggtcccgcg acctctacca cgtggccaac gccttcatcg ccgccggcga cgtggactcc 1080
qccaaccgct costcgacta cotogocaag gtggtgaagg acaacggcat gatcccgcag 1140
aacacctgga totooggcaa googtactgg acctocatoo agotogacga goaggcogac 1200
ecgateates tetectaceg ceteaagege tacgacetet acgaetecet egtgaageeg 1260
ctcgccgact teatcatcaa gatcggcccg aagaccggcc aggagcgctg ggaggagatc 1320
ggcqgctact ccccgccac gatggccgcc gaggtggccg gcctcacctg cgccgcctac 1380
attogeogage agaacaagga etacgagtee geecagaagt accaggagaa ggeegacaac 1440
tggcagaago toatogacaa octoacotao accgagaacg gcccgctcgg caacggccag 1500
tactacatec geategoogg ceteteogae ecgaacgeeg actteatgat caacategoe 1560
aacqqcqqcq qcqtqtacqa ccaqaaqqaq atcqtqqacc cqtccttcct cqaqctqqtq 1620
egectegged tgaagteege egacgaceeg aagateetea acaeeeteaa ggtggtggae 1680
tecaccatea aggtggacae ecegaaggge cegteetggt ategetacaa ecacgaegge 1740
tacggcgagc cgtccaagac cgagctgtac cacggcgccg gcaagggccg cctctggccg 1800
crectcaccy gegagegegg catglacgag ategeogecg geaaggaege caccecqtae 1860
gtgaaggcga tggagaagtt cgccaacgag ggcggcatca tctccgagca ggtgtgggag 1920
gacacegges tecegacega etecgeetes ecgeteaact gggeccaege egagtaegtg 1980
atcotottog cotocaacat cgagcacaag gtgotogaca tgccggacat cgtgtac
<210> 49
<211> 579
<212> PRT
<213> Rhizopus oryzae
<400> 49
Ala Ser Ile Pro Ser Ser Ala Ser Val Gln Leu Asp Ser Tyr Asn Tyr
```

```
10
Asp Gly Ser Thr Phe Ser Gly Lys Ile Tyr Val Lys Asn Ile Ala Tyr
                            25
Ser Lys Lys Val Thr Val Ile Tyr Ala Asp Gly Ser Asp Asn Trp Asn
                       40
      35
                                           45
Asn Asn Gly Asn Thr Ile Ala Ala Ser Tyr Ser Ala Pro Ile Ser Gly
                    55
Ser Asn Tyr Glu Tyr Trp Thr Phe Ser Ala Ser Ile Asn Gly Ile Lys
                . 70
                                    75
Glu Phe Tyr Ile Lys Tyr Glu Val Ser Gly Lys Thr Tyr Tyr Asp Asn
85 90 95
Asn Asn Ser Ala Asn Tyr Gln Val Ser Thr Ser Lys Pro Thr Thr 100 105 110
Thr Ala Thr Ala Thr Thr Thr Ala Pro Ser Thr Ser Thr Thr Thr 115 120 125
Pro Pro Ser Arg Ser Glu Pro Ala Thr Phe Pro Thr Gly Asn Ser Thr
130 140
Ile Ser Ser Trp Ile Lys Lys Gln Glu Gly Ile Ser Arg Phe Ala Met
145 150 155 160
Leu Arg Asn Ile Asn Pro Pro Gly Ser Ala Thr Gly Phe Ile Ala Ala
        165
                       170
Ser Leu Ser Thr Ala Gly Pro Asp Tyr Tyr Tyr Ala Trp Thr Arg Asp
180 185 190
Ala Ala Leu Thr Ser Asn Val Ile Val Tyr Glu Tyr Asn Thr Thr Leu
195 200 205
               200
Ser Gly Asn Lys Thr Ile Leu Asn Val Leu Lys Asp Tyr Val Thr Phe
                    215
                                         220
   210
Ser Val Lys Thr Gln Ser Thr Ser Thr Val Cys Asn Cys Leu Gly Glu
225 230 235 240
         230
Pro Lys Phe Asn Pro Asp Ala Ser Gly Tyr Thr Gly Ala Trp Gly Arg 245 250 255
Pro Gln Asn Asp Gly Pro Ala Glu Arg Ala Thr Thr Phe Ile Leu Phe
        260
                           265
Ala Asp Ser Tyr Leu Thr Gln Thr Lys Asp Ala Ser Tyr Val Thr Gly
   275 280 285
Thr Leu Lys Pro Ala Ile Phe Lys Asp Leu Asp Tyr Val Val Asn Val 290 295 300
Trp Ser Asn Gly Cys Phe Asp Leu Trp Glu Glu Val Asn Gly Val His 305 310 315 320
Phe Tyr Thr Leu Met Val Met Arg Lys Gly Leu Leu Leu Gly Ala Asp 325 330 335
Phe Ala Lys Arg Asn Gly Asp Ser Thr Arg Ala Ser Thr Tyr Ser Ser
                       345 350
Thr Ala Ser Thr Ile Ala Asn Lys Ile Ser Ser Phe Trp Val Ser Ser
                       360
                                           365
    355
Asn Asn Trp Ile Gln Val Ser Gln Ser Val Thr Gly Gly Val Ser Lys
370 380
Lys Gly Leu Asp Val Ser Thr Leu Leu Ala Ala Asn Leu Gly Ser Val
385 390 395 400
                   390
Asp Asp Gly Phe Phe Thr Pro Gly Ser Glu Lys Ile Leu Ala Thr Ala
                                410
                                                     415
             405
Val Ala Val Glu Asp Ser Phe Ala Ser Leu Tyr Pro Ile Asn Lys Asn
         420
                            425
Leu Pro Ser Tyr Leu Gly Asn Ser Ile Gly Arg Tyr Pro Glu Asp Thr
                 440
 435
Tyr Asn Gly Asn Gly Asn Ser Gln Gly Asn Ser Trp Phe Leu Ala Val
450 460
                    455
Thr Gly Tyr Ala Glu Leu Tyr Tyr Arg Ala Ile Lys Glu Trp Ile Gly
```

```
475
                    470
Asn Gly Gly Val Thr Val Ser Ser Ile Ser Leu Pro Phe Phe Lys Lys
                                                     495
                485
                                   490
Phe Asp Ser Ser Ala Thr Ser Gly Lys Lys Tyr Thr Val Gly Thr Ser
                                505
                                                    510
            500
Asp Phe Asn Asn Leu Ala Gln Asn Ile Ala Leu Ala Ala Asp Arg Phe
                            520
                                                525
       515
Leu Ser Thr Val Gln Leu His Ala His Asn Asn Gly Ser Leu Ala Glu
                        535
                                            540
    530
Glu Phe Asp Arg Thr Thr Gly Leu Ser Thr Gly Ala Arg Asp Leu Thr
                                     555
                  550
545
Trp Ser His Ala Ser Leu Ile Thr Ala Ser Tyr Ala Lys Ala Gly Ala
                                                        575
                                    570
                565
Pro Ala Ala
<210> 50
<211> 1737
<212> DNA
<213> Rhizopus oryzae
<400> 50
geotecated egtecteege eteegtgeag etegacteet acaactaega eggeteeace 60
ttctccggca aaatctacgt gaagaacatc gcctactcca agaaggtgac cgtgatctac 120
geogaegget eegaeaactg gaacaacaac ggeaacacca tegeogeete etacteegee 180
cogatetecg getecaacta egagtactgg acettetecg ectecateaa eggeateaag 240
gagttotaca toaagtacga ggtgtooggo aagacotact acgacaacaa caactoogoo 300
aactaccagg tgtccacctc caageegace accaceaceg ecacegeeac caceaceace 360
gooogtoca cotocarcae caccoogcog tecegeteeg ageoggecae ettecegace 420
ggcaactcca ccatctcctc ctggatcaag aagcaggagg gcatctcccg cttcgccatg 480
etecgeaca teaaceegee gggeteegee aceggettea tegeogeete cetetecace 540
geoggeoogg actactacta egeotogace egegacgeog coeteacete caacgtgate 600
grgracgagt acaacaccac cototoogge aacaagacca tootoaacgt gotoaaggac 660
tacqtqacct totocqtqaa gacccaqtoc acctccaccq tqtqcaactq cotcqqcqaq 720
ccgaagttca acceggacge etceggetac accggegeet gggggeegeec geagaacgae 780
ggcccggccg agcgcgccac cacettcate etettegecg actectacet cacecagace 840
aaggacgeet cetacgtgae eggeaceete aageeggeea tetteaagga eetegaetae 900
gragraacq raragracaa cagcractro gaccroraga aggagargaa cagcaraca 960
ttotacacco toatggtgat gogcaagggo otcotoctog gogcogactt ogccaagego 1020
aacggcgact ccacccgcgc ctccacctac tcctccaccg cctccaccat cgccaacaaa 1080
atotoctoot totgggtgto otocaacaac tggatacagg tgtoccagto cgtgaccggo 1140
gocgtateca agaagageet egacatatee acceteeteg eegecaacet egacteegta 1200
gacgacgget tetteacece gggeteegag aagateeteg eeacegeegt ggeegtggag 1260
gactocticg ecteseteta eccgateaac aagaacetee egtectacet eggeaactee 1320
atoggoogot accoggagga cacctacaac ggcaacggca actoccaggg caactootgg 1380
trectegoog tgacoggota ogcogagotg tactacogog coatcaagga gtggatoggo 1440
aacgocggcg tgaccgtgtc ctccatctcc ctcccgttct tcaagaagtt cgactcctcc 1500
gocacctocg gcaagaagta caccgtgggc acctccgact tcaacaacct cgcccagaac 1560
ategeceteg eegeegaceg ettectetee acegtgeage tecaegeeca caacaaegge 1620
tecetegeeg aggagttega eegeaceace ggeeteteea eeggegeeeg egaceteace 1680
tggtcccacg cotcootcat caccgcotco tacgccaagg coggegeecc ggccgcc
<210> 51
<211> 439
<212> PRT
```

Page 54

<213> Artificial Sequence

<220> <223> synthetic <400> 51 Met Ala Lys His Leu Ala Ala Met Cys Trp Cys Ser Leu Leu Val Leu 10 1 Val Leu Leu Cys Leu Gly Ser Gln Leu Ala Gln Ser Gln Val Leu Phe 25 Gln Gly Phe Asn Trp Glu Ser Trp Lys Lys Gln Gly Gly Trp Tyr Asn 35 40 Tyr Leu Leu Gly Arg Val Asp Asp Ile Ala Ala Thr Gly Ala Thr His 55 Val Trp Leu Pro Gln Pro Ser His Ser Val Ala Pro Gln Gly Tyr Met 65 70 75 Pro Gly Arg Leu Tyr Asp Leu Asp Ala Ser Lys Tyr Gly Thr His Ala 8.5 90 Glu Leu Lys Ser Leu Thr Ala Ala Phe His Ala Lys Gly Val Gln Cys 100 105 110 Val Ala Asp Val Val Ile Asn His Arg Cys Ala Asp Tyr Lys Asp Gly 115 120 125 Arg Gly Ile Tyr Cys Val Phe Glu Gly Gly Thr Pro Asp Ser Arg Leu 130 135 Asp Trp Gly Pro Asp Met Ile Cys Ser Asp Asp Thr Gln Tyr Ser Asn 150 155 145 Gly Arg Gly His Arg Asp Thr Gly Ala Asp Phe Ala Ala Ala Pro Asp 170 175 165 Ile Asp His Leu Asn Pro Arg Val Gln Gln Glu Leu Ser Asp Trp Leu 190 180 185 Asn Trp Leu Lys Ser Asp Leu Gly Phe Asp Gly Trp Arg Leu Asp Phe 195 200 205 Ala Lys Gly Tyr Ser Ala Ala Val Ala Lys Val Tyr Val Asp Ser Thr 210 220 215 Ala Pro Thr Phe Val Val Ala Glu Ile Trp Ser Ser Leu His Tyr Asp 230 235 Gly Asn Gly Glu Pro Ser Ser Asn Gln Asp Ala Asp Arg Gln Glu Leu 250 255 245 Val Asn Trp Ala Gln Ala Val Gly Gly Pro Ala Ala Ala Phe Asp Phe 260 265 270 260 Thr Thr Lys Gly Val Leu Gln Ala Ala Val Gln Gly Glu Leu Trp Arg 275 280 285 Met Lys Asp Gly Asn Gly Lys Ala Pro Gly Met Ile Gly Trp Leu Pro 290 295 300 Glu Lys Ala Val Thr Phe Val Asp Asn His Asp Thr Gly Ser Thr Gln 315 310 Asn Ser Trp Pro Phe Pro Ser Asp Lys Val Met Gln Gly Tyr Ala Tyr 330 335 325 He Leu Thr His Pro Gly Thr Pro Cys Ile Phe Tyr Asp His Val Phe 340 345 350Asp Trp Asn Leu Lys Gln Glu Ile Ser Ala Leu Ser Ala Val Arg Ser 365 355 360 Arg Asn Gly Ile His Pro Gly Ser Glu Leu Asn Ile Leu Ala Ala Asp 375 380 Gly Asp Leu Tyr Val Ala Lys Ile Asp Asp Lys Val Ile Val Lys Ile 385 390 395 400 Gly Ser Arg Tyr Asp Val Gly Asn Leu Ile Pro Ser Asp Phe His Ala 410 415 405 Val Ala His Gly Asn Asn Tyr Cys Val Trp Glu Lys His Gly Leu Arg 425 430 420

```
Val Pro Ala Gly Arg His His
        435
<210> 52
<211> 1320
<212> DNA
<213> Artificial Sequence
<220> -
<223> synthetic
<400> 52
atggcgaage acttggctge catgtgctgg tgcageetee tagtgcttgt actgctctge 60
traggeteec agergacea atcocaggie eteticage getteaacte geagteging 120
aagaagcaag gtgggtggta caactacete etggggeggg tggaegacat egeegegaeg 180
ggggccacgc acgtetgget eccgcagecg tegeactegg tggegeegea ggggtacatg 240
cccqqccqqc tetacqacet ggacqcqtcc aagtacqqca cccacqcqqa qctcaaqtcq 300
ctcaccgcgg cgttccacgc caagggcgtc cagtgcgtcg ccgacgtcgt gatcaaccac 360
cgctgcgccg actacaagga cggccgcggc atctactgcg tcttcgaggg cggcacgccc 420
gacageegee tegactgggg eccegacatg atetgeageg acgacaegea gtactecaae 480
gggcgcgggc accgcgacac gggggccgac ttcgccgccg cgcccgacat cgaccacctc 540
aacccqcqcq tqcaqcaqqa qctctcqqac tqqctcaact qqctcaaqtc cqacctcqqc 600
ttogacogot ggogootoga ottogocaag ggotactoog cogoogtogo caaggtgtac 660
gregacagea ecqueeceae ettegregre geogagatat ggageteeet ecactaegae 720
ggcaacggcg agccgtccag caaccaggac gccgacaggc aggagctggt caactgggcg 780
caggeggtgg geggeeege egeggegtte gaetteacea ceaagggegt getgeaggeg 840
qccqtccaqq qcqaqctqtq qcqcatqaaq qacqqcaacq qcaaqqcqcc cqqqatqatc 900
ggctggctgc cggagaaggc cgtcacgttc gtcgacaacc acgacacegg ctccacgcag 960
aactogtggo cattococto ogacaaggto atgoagggot acgeotatat cotoacgoac 1020
ccaggaacte catgeatett etacgaceae gttttegaet ggaacetgaa geaggagate 1080
agegegetgt etgeggtgag gteaagaaac gggateeace eggggagega getgaacate 1140
ctcgccgccg acogggatet ctacgtcgcc aagattgacg acaaggtcat cgtgaagate 1200
quateacqqt acqacqtcqq gaacctqate coctcagact tecacqccqt tgcccctggc 1260
aacaactact qcqtttqqqa qaaqcacqqt ctqaqaqttc caqcqqqqcq qcaccactag 1320
<210> 53
<211> 45
<212> PRT
<213> Artificial Sequence
<220>
<223> synthetic
<400> 53
Ala Thr Gly Gly Thr Thr Thr Ala Thr Thr Thr Gly Ser Gly Gly
                5
                                    10
)
Val Thr Ser Thr Ser Lys Thr Thr Thr Ala Ser Lys Thr Ser Thr
           20
                                25
Thr Thr Ser Ser Thr Ser Cys Thr Thr Pro Thr Ala Val
<210> 54
<211> 137
<212> DNA
<213> Artificial Sequence
```

Page 56

```
<220>
<223> synthetic
<400> 54
qccaccqqcq qcaccaccac caccqccacc accaccqqct ccqqcqqcqt qacctccacc 60
tecaagacea ceaceacege etecaagace tecaecacea ectectecae etectgeace 120
accccgaccg ccgtgtc
<210> 55
<211> 300
<212> PRT
<213> Pyrococcus furiosus
<400> 55
Ile Tyr Phe Val Glu Lys Tyr His Thr Ser Glu Asp Lys Ser Thr Ser
                                 10
                                                    15
Asn Thr Ser Ser Thr Pro Pro Gln Thr Thr Leu Ser Thr Thr Lys Val
          20
                            25
                                               30
Leu Lys Ile Arg Tyr Pro Asp Asp Gly Glu Trp Pro Gly Ala Pro Ile
     35
                         40
                                           45
Asp Lys Asp Gly Asp Gly Asn Pro Glu Phe Tyr Ile Glu Ile Asn Leu
                      55
Trp Asn Ile Leu Asn Ala Thr Gly Phe Ala Glu Met Thr Tyr Asn Leu
               70
                                     75
Thr Ser Gly Val Leu His Tyr Val Gln Gln Leu Asp Asn Ile Val Leu
                               90
             8.5
Arg Asp Arg Ser Asn Trp Val His Gly Tyr Pro Glu Ile Phe Tyr Gly
          100
                           105
                                               110
Asn Lys Pro Trp Asn Ala Asn Tyr Ala Thr Asp Gly Pro Ile Pro Leu
                         120
                                         125
      115
Pro Ser Lys Val Ser Asn Leu Thr Asp Phe Tyr Leu Thr Ile Ser Tyr
                                       140
                     135
  130
Lys Leu Glu Pro Lys Asn Gly Leu Pro Ile Asn Phe Ala Ile Glu Ser
                  150
                                   155
                                                     160
Trp Leu Thr Arg Glu Ala Trp Arg Thr Thr Gly Ile Asn Ser Asp Glu
                              170
             165
                                                  175
Gln Glu Val Met Ile Trp Ile Tyr Tyr Asp Gly Leu Gln Pro Ala Gly
          180
                             185
                                              190
Ser Lys Val Lys Glu Ile Val Val Pro Ile Ile Val Asn Gly Thr Pro
                         200
                                            205
Val Asn Ala Thr Phe Glu Val Trp Lys Ala Asn Ile Gly Trp Glu Tyr
 210
                      215
                                        220
Val Ala Phe Arg Ile Lys Thr Pro Ile Lys Glu Gly Thr Val Thr Ile
                                   235
        230
225
Pro Tyr Gly Ala Phe Ile Ser Val Ala Ala Asn Ile Ser Ser Leu Pro
           245
                              250
Asn Tyr Thr Glu Leu Tyr Leu Glu Asp Val Glu Ile Gly Thr Glu Phe
                            265
           260
Gly Thr Pro Ser Thr Thr Ser Ala His Leu Glu Trp Trp Ile Thr Asn
                  280
      275
He Thr Leu Thr Pro Leu Asp Arg Pro Leu Ile Ser
                      295
                                        300
<210> 56
<211> 903
<212> DNA
<213> Pyrococcus furiosus
```

Page 57

```
<400> 56
atotacticg tggagaagta ccacacctoc gaggacaagt ccacctccaa cacctcctcc 60
accoegoege agaceacet etecaceace aaggtgetea agateegeta eceggacgae 120
ggcgagtggc ccggcgccc gatcgacaag gacggcgacg gcaacccgga gttctacatc 180
gagatcaacc totggaacat cotcaacgoo accggottog cogagatgac ctacaacoto 240
actagtogog tgotocacta cgtgcagcag ctcgacaaca tcgtgctccg cgaccgctcc 300
aactgggtgc acggctaccc ggaaatette tacggcaaca agccgtggaa cgccaactac 360
gccaccqacg gcccgatccc gctcccgtcc aaggtgtcca acctcaccga cttctacctc 420
accateteet acaagetega geogaagaac ggteteeega teaacttege categagtee 480.
tggctcaccc gcgaggcctg gcgcaccacc ggcatcaact ccgacgagca ggaggtgatg 540
atorggator actacgacgg cotocageoc gegggeteca aggregaagga gategrage 600
ccgatcatcg tgaacggcac cccggtgaac gccaccttcg aggtgtggaa ggccaacatc 660
ggctgggagt acgtggcctt ccgcatcaag accccgatca aggagggcac cgtgaccatc 720
ccgtacggcg ccttcatctc cgtggccgcc aacateteet ecetecegaa etacaecgag 780
aagtacctcg aggacgtgga gatcggcacc gagttcggca ccccgtccac cacctccgcc 840
cacctcgagt ggtggatcac caacatcacc ctcaccccgc tcgaccgccc gctcatctcc 900
taq
<210> 57
<211> 387
<212> PRT
<213> Thermus flavus
<400> 57
Met Tyr Glu Pro Lys Pro Glu His Arg Phe Thr Phe Gly Leu Trp Thr
                                   10
1
Val Asp Asn Val Asp Arg Asp Pro Phe Gly Asp Thr Val Arg Glu Arg
           20
Leu Asp Pro Val Tyr Val Val His Lys Leu Ala Glu Leu Gly Ala Tyr
                                              45
                           40
       35
Gly Val Asn Leu His Asp Glu Asp Leu Ile Pro Arg Gly Thr Pro Pro
                       5.5
   50
Gin Glu Arg Asp Gln Ile Val Arg Arg Phe Lys Lys Ala Leu Asp Glu
                                       75
                   70
Thr Val Leu Lys Val Pro Met Val Thr Ala Asn Leu Phe Ser Glu Pro
                                 90
               85
Ala Phe Arg Asp Gly Ala Ser Thr Thr Arg Asp Pro Trp Val Trp Ala
                                                  110
                              105
Tyr Ala Leu Arg Lys Ser Leu Glu Thr Met Asp Leu Gly Ala Glu Leu
                                               125
      115
                           120
Gly Ala Glu Ile Tyr Met Phe Trp Met Val Arg Glu Arg Ser Glu Val
                      135
   130
Glu Ser Thr Asp Lys Thr Arg Lys Val Trp Asp Trp Val Arg Glu Thr
                                       155
                   150
Leu Asn Phe Met Thr Ala Tyr Thr Glu Asp Gln Gly Tyr Gly Tyr Arg
                                                       175
                                  170
               165
Phe Ser Val Glu Pro Lys Pro Asn Glu Pro Arg Gly Asp Ile Tyr Phe
                                                   190
                              185
           130
Thr Thr Val Gly Ser Met Leu Ala Leu Ile His Thr Leu Asp Arg Pro
                                               205
       195
                           200
Glu Arg Phe Gly Leu Asn Pro Glu Phe Ala His Glu Thr Met Ala Gly
                       215
                                           220
Leu Asn Phe Asp His Ala Val Ala Gln Ala Val Asp Ala Gly Lys Leu
                                     235
225
                   230
Phe His Ile Asp Leu Asn Asp Gln Arg Met Ser Arg Phe Asp Gln Asp
                                  250
                                                       255
              245
Leu Arg Phe Gly Ser Glu Asn Leu Lys Ala Gly Phe Phe Leu Val Asp
                               265
           260
```

```
Leu Leu Glu Ser Ser Gly Tyr Gln Gly Pro Arg His Phe Glu Ala His
                            280
        275
Ala Leu Arg Thr Glu Asp Glu Glu Gly Val Trp Thr Phe Val Arg Val
                                            300
    290
                        295
Cys Met Arg Thr Tyr Leu Ile Ile Lys Val Arg Ala Glu Thr Phe Arg
305
                    310
                                        315
Glu Asp Pro Glu Val Lys Glu Leu Leu Ala Ala Tyr Tyr Gln Glu Asp
                                                        335
                325
                                    330
Pro Ala Thr Leu Ala Leu Leu Asp Pro Tyr Ser Arg Glu Lys Ala Glu
                                                   350
                                345
Ala Leu Lys Arg Ala Glu Leu Pro Leu Glu Thr Lys Arg Arg Arg Gly
                                                365
                            360
Tyr Ala Leu Glu Arg Leu Asp Gln Leu Ala Val Glu Tyr Leu Leu Gly
    370
                        375
                                            380
Val Arg Gly
<210> 58
<211> 978
<212> DNA
<213> Artificial Sequence
<220>
<223> synthetic
<400> 58
atggggaaga acggcaacct gtgctgcttc tctctgctgc tgcttcttct cgccgggttg 60
gegreeggee areaaateta ettegrggag aagtaceaca eeteegagga caagteeace 120
tccaacact cetecacee geogragace acceteteca ccaccaaggt geteaagate 180
egetaceegy acgaeggtga gtggeeegge geeeegateg acaaggaegg egaeggeaac 240
coggagitet acatogagat caacetotgg aacateetca acgecacegg ettegeegag 300
atgacctaca accteactag tggcgtgete cactacgtge ageagetega caacategtg 360
ctccgcgacc gctccaactg ggtgcacggc tacccggaaa tcttctacgg caacaagccg 420
tggaacgcca actacgccac cgacggcccg atcccgctcc cgtccaaggt gtccaacctc 480
accgactict acctcaccat cicctacaag cicgageega agaaeggiet ecegateaac 540
ttcgccatcg agtcctggct cacccgcgag gcctggcgca ccaccggcat caactccgac 600
gagcaggagg tgatgatetg gatetactae gaeggeetee agecegeggg etecaaggtg 660
aaggagateg togtgeegat categtgaae ggeaceegg tgaaegeeae ettegaggtg 720
tggaaggcca acateggetg ggagtacgtg gcetteegea teaagaeeee gateaaggag 780
ggcaccgtga ccatcccgta cggcgccttc atctccgtgg ccgccaacat ctcctccctc 840
ccgaactaca ccgagaagta cctcgaggac gtggagatcg gcaccgagtt cggcaccccg 900
tecaccacet eggeeacet egagtggtgg atcaccaaca teacceteae ecegetegae 960
                                                                  978
cgcccgctca tctcctag
<210> 59
<211> 1920
<212> DNA
<213> Aspergillus niger
<400> 59
atgreettee geteceteet egecetetee ggeetegtgt geaceggeet egecaaegtg 60
atotocaago gogocaccot ogactootgg ototocaacg aggocaccgt ggcccgcacc 120
gccatcctca acaacategg egeegaegge geetgggtgt eeggegeega eteeggeate 180
gtggtggcct ccccgtccac cgacaacccg gactacttct acacctggac ccgcgactcc 240
ggcctcgtgc tcaagaccet cgtggaccte tteegeaacg gegacacete cetectetee 300
accategaga actacatete egeceaggee ategtgeagg geateteeaa eeegteegge 360
gacctetet eeggegeegg ceteggegag eegaagttea aegtggaega gacegeetae 420
```

accogetect ggggeegeec geagegegae ggeeeggeec teegegeeae egecatgate 480

```
ggetteggee agtggeteet egacaaegge tacaceteea eegecaeega categtgtgg 540
cogctogtgo gcaacgacot otoctacgtg gcccagtact ggaaccagac cggctacgac 600
ctctgggagg aggtgaacgg ctcctccttc ttcaccatcg ccgtgcagca ccgcgccctc 660
gtggaggget eegeettege caeegeegtg ggeteeteet geteetggtg egaeteecag 720
geocoggaga tectetgeta ectecagtee ttetggaceg geteetteat cetegecaae 780
ttegacteet ecegeteegg caaggaegee aacacette teggeteeat ceacacette 840
gacccggagg ccgcctgcga cgactccacc ttccagccgt gctccccgcg cgccctcgcc 900
aaccacaagg aggtggtgga ctccttccgc tccatctaca ccctcaacga cggcctctcc 960
gactccgagg ccgtggccgt gggccgctac ccggaggaca cctactacaa cggcaacccg 1020
tggtteetet geaccetege egeegeegag eagetetaeg aegeeeteta ecagtgggae 1080
aagcaggget ceetegaggt gacegaegtg teeetegact tetteaagge cetetactee 1140
gacgoogoca coggoacota etectoctoc tectecacet actoetecat egtggacgee 1200
greaagacet tegeogaegg ettegtgtee ategtggaga eccaegeege etecaacgge 1260
tocatgleeg ageaglacga caagloogae ggogagoago toloogooog egacoloaco 1320
togtcctacg cogcoctect cacegocaac aacegocgca acteogtggt geoggeetee 1380
tgggggaga cotocgecto etcegtgeeg ggeacetgeg eegecacete egecategge 1440
acctactact cogtgacogt gacotootgg cogtocatog tggccacogg cggcaccace 1500
accaccacca coccaaccag eteogactor ataacctora cotocaagac caccaccacc 1560
geotecaaga ectecacete cacetectee acetectgea ceacecegae egeogtggee 1620
gtqacetteg aceteacege caccaceae tacggegaga acatetacet egtgggetee 1680
atctcccage teggegactg ggagacetee gacggcateg eceteteege egacaagtae 1740
acctectecy acceptety gtacgtgace gtgaccetee eggeeggega gteettegag 1800
tacaagttca teegeatega gteegaegae teegtggagt gggagteega eeegaacege 1860
gagradaceg tgoogdagge otgoggdace tedacegeda cegtgacega cacetggege 1920
<210> 60
<211> 6
<212> PRT
<213> Artificial Sequence
<220>
<223> synthetic
<400> 60
Ser Glu Lys Asp Glu Leu
<210> 61
<211> 561
<212> DNA
<213> Artificial Sequence
<220>
<223> Xylanase BD7436
<220>
<221> CDS
<222> (1)..(561)
<400> 61
atg get age ace the tac tgg eat the tgg ace gae gge ate gge ace
                                                                      48
Met Ala Ser Thr Phe Tyr Trp His Leu Trp Thr Asp Gly Ile Gly Thr
```

gtg aac get acc aac ggc agc gac ggc aac tac agc gtg agc tgg agc Val Asn Ala Thr Asn Gly Ser Asp Gly Asn Tyr Ser Val Ser Trp Ser 20 25	96
aac tgc ggc aac ttc gtg gtg ggc aag ggc tgg acc acc ggc agc gct Asn Cys Gly Asn Phe Val Val Gly Lys Gly Trp Thr Thr Gly Ser Ala 35 40 45	144
acc agg gtg atc aac tac aac gct cat gct ttc agc gtg gtg ggc aac Thr Arg Val Ile Asn Tyr Asn Ala His Ala Phe Ser Val Val Gly Asn 50 55 60	192
get tac ttg get ttg tac gge tgg ace agg aac age ttg ate gag tac Ala Tyr Leu Ala Leu Tyr Gly Trp Thr Arg Asn Ser Leu Ile Glu Tyr 65 70 75 80	240
tac gtg gtg gac agc tgg ggc acc tac agg cca acc ggc acc tac aag Tyr Val Val Asp Ser Trp Gly Thr Tyr Arg Pro Thr Gly Thr Tyr Lys 85 90 95	288
ggc acc gtg acc agc gac ggc ggc acc tac gac atc tac acc acc Gly Thr Val Thr Ser Asp Gly Gly Thr Tyr Asp Ile Tyr Thr Thr 100 105 110	336
agg acc aac gct cca agc atc gac ggc aac aac acc acc ttc acc caa Arg Thr Asn Ala Pro Ser Ile Asp Gly Asn Asn Thr Thr Phe Thr Gln 115 120 125	384
ttc tgg agc gtg agg caa agc aag agg cca atc ggc acc aac acc Phe Trp Ser Val Arg Gln Ser Lys Arg Pro Ile Gly Thr Asn Asn Thr 130 140	432
atc acc ttc age aac cat gtg aac gct tgg aag age aag ggc atg aac Ile Thr Phe Ser Asn His Val Asn Ala Trp Lys Ser Lys Gly Met Asn 145 150 155 160	480
ttg ggc agc agc tgg agc tac caa gtg ttg gct acc gag ggc tac caa Leu Gly Ser Ser Trp Ser Tyr Gln Val Leu Ala Thr Glu Gly Tyr Gln 165 170	528
age age gge tae age aac gtg ace gtg tgg tag Ser Ser Gly Tyr Ser Asn Val Thr Val Trp 180 185	561
<210> 62 <211> 186 <212> PRT <213> Artificial Sequence	
<220> <223> Synthetic Construct	
<400> 62	
Met Ala Ser Thr Phe Tyr Trp His Leu Trp Thr Asp Gly Ile Gly Thr 10 15	

Page 61

Val Asn Ala Thr Asn Gly Ser Asp Gly Asn Tyr Ser Val Ser Trp Ser

Asn Cys Gly Asn Phe Val Val Gly Lys Gly Trp Thr Thr Gly Ser Ala

Thr Arg Val Ile Asn Tyr Asn Ala His Ala Phe Ser Val Val Gly Asn

Ala Tyr Leu Ala Leu Tyr Gly Trp Thr Arg Asn Ser Leu Ile Glu Tyr 65 70 80

Tyr Val Val Asp Ser Trp Gly Thr Tyr Arg Pro Thr Gly Thr Tyr Lys

Gly Thr Val Thr Ser Asp Gly Gly Thr Tyr Asp Ile Tyr Thr Thr

Arg Thr Asn Ala Pro Ser lie Asp Gly Asn Asn Thr Thr Phe Thr Gln

Phe Trp Ser Val Arg Gln Ser Lys Arg Pro Ile Gly Thr Asn Asn Thr 130

Ile Thr Phe Ser Asn His Val Asn Ala Trp Lys Ser Lys Gly Met Asn 145 150

Leu Gly Ser Ser Trp Ser Tyr Gln Val Leu Ala Thr Glu Gly Tyr Gln 170 165

Ser Ser Gly Tyr Ser Asn Val Thr Val Trp

<210> 63

<211> 561
<212> DNA
<213> Artificial Sequence

<220>

<223> Xylanase BD6002A

<220>

<221> CDS <222> (1)..(561)

Page 62

	•												-				
	gct	63 agc Ser															48
		gct Ala															96
		ggc Gly 35														1	4 4
		gtg Val														1	92
ggc Gly 65	tac Tyr	t t g Leu	gct Ala	ttg Leu	tac Tyr 70	ggc Gly	tgg Trp	acc Thr	agg Arg	aac Asn 75	agc Ser	ttg Leu	atc Ile	gag Glu	tac Tyr 80	2	40
		gtg Val														21	88
		gtg Val														3	36
		aac Asn 115														31	8 4
ttc Phe	tgg Trp 130	agc Se <i>r</i>	gtg Val	agg Arg	caa Gln	agc Ser 135	aag Lys	agg Arg	cca Pro	atc Ile	ggc Gly 140	acc Thr	aac Asn	aac Asn	acc Thr	4.	32
atc Ile 145	acc Thr	ttc Phe	agc Ser	aac Asn	cat His 150	gtg Val	aac Asn	gct Ala	tgg Trp	aag Lys 155	agc Ser	aag Lys	ggc Gly	atg Met	aac Asn 160	4 8	80
ttg Leu	Gly	agc Ser	agc Ser	tgg Trp 165	agc Ser	tac Tyr	caa Gln	gtg Val	ttg Leu 170	gct Ala	acc Thr	gag Glu	ggc Gly	tac Tyr 175	caa Gln	52	28
		ggc Gly								tag						5 6	51
<210 <211 <212 <213	> 1 !> P	4 86 RT rtif	ıcia	l Se	quen	ce											
<220 <223		ynth	etic	Con	stru	Ct											

<400> 64

Met Ala Ser Thr Asp Tyr Trp Gln Asn Trp Thr Asp Gly Gly Gly Thr

Val Asn Ala Thr Asn Gly Ser Asp Gly Asn Tyr Ser Val Ser Trp Ser

Asn Cys Gly Asn Phe Val Val Gly Lys Gly Trp Thr Thr Gly Ser Ala

Thr Arg Val Ile Asn Tyr Asn Ala Gly Ala Phe Ser Pro Ser Gly Asn

Gly Tyr Leu Ala Leu Tyr Gly Trp Thr Arg Asn Ser Leu Ile Glu Tyr 65 70 75 80

Tyr Val Val Asp Ser Trp Gly Thr Tyr Arg Pro Thr Gly Thr Tyr Lys 85 90 95

Gly Thr Val Thr Ser Asp Gly Gly Thr Tyr Asp Ile Tyr Thr Thr Thr

Arg Thr Asn Ala Pro Ser Ile Asp Gly Asn Asn Thr Thr Phe Thr Gln

Phe Trp Ser Val Arg Gln Ser Lys Arg Pro Ile Gly Thr Asn Asn Thr

The Thr Phe Ser Ash His Val Ash Ala Trp Lys Ser Lys Gly Met Ash

Leu Gly Ser Ser Trp Ser Tyr Gln Val Leu Ala Thr Glu Gly Tyr Gln 165 170 175 165

Ser Ser Gly Tyr Ser Asn Val Thr Val Trp

<210> 65 <211> 561 <212> DNA <213> Artificial Sequence

<220>

<223> Xylanase BD6002B

Page 64

	CDS (1)(56	1)										
<400> 6 atg gcc Met Ala 1												48
gtg aac Val Asn												96
aac tgc Asn Cys			. Val G									144
acc cgc Thr Arg 50												192
ggc tac Gly Tyr 65	ctc gcc Leu Ala	ctc tac Leu Ty: 70	gçc t Gly T	gg acc rp Thr	cgc Arg	aac Asn 75	tcc Ser	ctc Leu	atc Ile	gag Glu	tac Tyr 80	240
tac gtg	gtg gac Val Asp	tcc tgc Ser Trp 85	g ggc a o Gly T	cc tac hr Tyr	ege Arg 90	ccg Pro	acc Thr	ggc Gly	acc Thr	tac Tyr 95	aag Lys	288
ggc acc Gly Thr												336
ege ace a			Ile A									384
ttc tgg Phe Trp : 130	tcc gtg Ser Val	cgc cac Arg Glr	tec a Ser L 135	ag cgc ys Arg	ccg Pro	atc Ile	ggc Gly 140	acc Thr	aac Asn	aac Asn	acc Thr	432
atc acc Ile Thr 145			Val A									480
ctc ggc (Leu Gly S	tcc tcc Ser Ser	tgg tcd Trp Sei 165	tac c Tyr G	ag gtg ln Val	ctc Leu 170	gcc Ala	acc Thr	gag Glu	ggc Gly	tac Tyr 175	cag Gln	528
tcc tcc (Ser Ser (ggc tac Gly Tyr 180	toc aad Ser Asr	gtg a Val T	cc gtg hr Val 185	tgg Trp	tga						561

<210> 66 <211> 186

Page 65

<212> PRT <213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 66

Met Ala Ser Thr Asp Tyr Trp Gln Asn Trp Thr Asp Gly Gly Thr

Val Asn Ala Thr Asn Gly Ser Asp Gly Asn Tyr Ser Val Ser Trp Ser

Asn Cys Gly Asn Phe Val Val Gly Lys Gly Trp Thr Thr Gly Ser Ala

Thr Arg Val Ile Asn Tyr Asn Ala Gly Ala Phe Ser Pro Ser Gly Asn

Gly Tyr Leu Ala Leu Tyr Gly Trp Thr Arg Asn Ser Leu Ile Glu Tyr

Tyr Val Val Asp Ser Trp Gly Thr Tyr Arg Pro Thr Gly Thr Tyr Lys

Gly Thr Val Thr Ser Asp Gly Gly Thr Tyr Asp Ile Tyr Thr Thr Thr

Arg Thr Asn Ala Pro Ser Ile Asp Gly Asn Asn Thr Thr Phe Thr Gln 115

Phe Trp Ser Val Arg Gln Ser Lys Arg Pro Ile Gly Thr Asn Asn Thr 130

lle Thr Phe Ser Asn His Val Asn Ala Trp Lys Ser Lys Gly Met Asn 150 155

Leu Gly Ser Ser Trp Ser Tyr Gln Val Leu Ala Thr Glu Gly Tyr Gln

Ser Ser Gly Tyr Ser Asn Val Thr Val Trp 185

<210> 67 <211> 2071

Page 66

<212> DNA <213> Oryza sativa <220> <221> misc_feature <222> (1)..(2071) <223> Promoter <400> 67 tecatgoigt cotactactt gottcatccc ottotacatt tigitotggt titiggootg 60 120 cattttcggat catgatgtat gtgatttcca atctgctgca atatgaatgg agactctgtg 180 ctaaccatca acaacatgaa atgettatga ggeetttget gageageeaa tettgeetgt gittalgict tcacaggccg aattcctctg tittgttttt caccitcaat atttggaaac 240 attitationag grigititgig todaggoota taaatoatac atgatgitgi ogtatiggat 300 gtgaatgtgg tggcgtgttc agtgccttgg atttgagttt gatgagagtt gcttctgggt 360 caccactoac cattategat geteetette ageataaggt aaaagtette eetgtttaeg 420 480 ttatttttacc cactatggtt gcttgggttg gttttttcct gattgcttat gccatggaaa qtcatttqat atqttqaact tgaattaact gtagaattgt atacatgttc catttgtgtt 540 gtacttcctt cttttctatt agtagcctca gatgagtgtg aaaaaaacag attatataac 600 ttqccctata aatcatttqa aaaaaatatt gtacagtgag aaattgatat atagtgaatt 660 tttaaqaqca tgttttccta aagaagtata tattttctat gtacaaaggc cattgaagta 720 attgtagata caggataatg tagacttttt ggacttacac tgctaccttt aagtaacaat 780 catgagcaat agtgttgcaa tgatatttag gctgcattcg tttactctct tgatttccat 840 gagcacgett eccaaactgt taaactetgt gttttttgee aaaaaaaaat geataggaaa 900 qttqctttta aaaaatcata tcaatccatt ttttaagtta tagctaatac ttaattaatc 960 1020 atgogotiant augitoactor gittiticqia ciagagagat tgittigaac cagcactoaa qaacacacc ttaacccaqc caaataatgc tacaacctac cagtccacac ctcttgtaaa 1080 gcatttgttg catggaaaag ctaagatgac agcaacctgt tcaggaaaac aactgacaag 1140 1200 ctcatagoga gagggagett ttggaaaggt geegtgeagt teaaacaatt agttageagt agggtettgg tttttgctca cagcaataag aagttaatca tggtgtaggc aacccaaata 1260 1320 aaacaccaaa atatgcacaa ggcagtttgt tgtattctgt agtacagaca aaactaaaag taatgaaaga agatgtggtg ttagaaaagg aaacaatatc atgagtaatg tgtgggcatt 1380 atgggaccac gaaataaaaa gaacattttg atgagtcgtg tatcctcgat gagcctcaaa 1440

```
agttototoa cocoggataa gaaaccotta agcaatgtgo aaagtttgoa ttotocactg
                                                                    1500
                                                                    1560
acataatqca aaataaqata tcatcgatga catagcaact catgcatcat atcatgcctc
totcaacota ttoattocta otcatotaca taagtatott cagotaaatg ttagaacata
                                                                    1620
aacccataag tcacgtttga tgagtattag gcgtgacaca tgacaaatca cagactcaag
                                                                    1680
caagataaag caaaatgatg tgtacataaa actccagagc tatatgtcat attgcaaaaa
                                                                    1740
gaggagaget tataagacaa ggeatgaete acaaaaatte atttgeettt egtgteaaaa
                                                                    1800
agaggagggc tttacattat ccatgtcata ttgcaaaaga aagagagaaa gaacaacaca
                                                                    1860
atgotgogto aattatacat atotgtatgt coatcattat toatcoacct ttogtgtaco
                                                                    1920
                                                                    1980
acacttcata tatcatgagt cacttcatgt ctggacatta acaaactcta tcttaacatt
                                                                    2040
tagatocaag agootttato toactataaa tgoacgatga tttotcattg tttotcacaa
                                                                    2071
aaagcattca gttcattagt cctacaacaa c
```

```
<210> 68
```

<220>

<221> SIGNAL

<222> (1). (79) <223> Maize waxy signal sequence.

<400> 68

Met Leu Ala Ala Leu Ala Thr Ser Gln Leu Val Ala Thr Arg Ala Gly

Leu Gly Val Pro Asp Ala Ser Thr Phe Arg Arg Gly Ala Ala Gln Gly

Leu Arg Gly Ala Arg Ala Ser Ala Ala Ala Asp Thr Leu Ser Met Arg

Thr Ser Ala Arg Ala Ala Pro Arg His Gln His Gln Gln Ala Arg Arg 50 55

Gly Ala Arg Phe Pro Ser Leu Val Val Cys Ala Ser Ala Gly Ala

<210> 69 <211> 1005

Page 68

<211> 79

<212> PRT

<213> Zea mays

<21 <21	_	DNA Arti	fici	al S	eque	nce										
<22 <22	_	Synt	het i	c Br	omel	ain .	Segu	ence								
<22 <22 <22	i > (2 >	CDS (1). Synt!			omel	ain										
<40 atg Met 1	gcc	69 tgg Trp	aag Lys	gtg Val 5	cag Gln	gtg Val	gtg Val	ttc Phe	ctc Leu 10	ttc Phe	ctc Leu	ttc Phe	ctc Leu	tgc Cys 15	gtg Val	48
						gcc Ala										96
atg Met	atg Met	aag Lys 35	cgc Arg	ttc Phe	gag Glu	gag Glu	tgg Trp 40	atg Met	gtg Val	gag Glu	tac Tyr	ggc Gly 45	cgc Arg	gtg Val	tac Tyr	144
aag Lys	gac Asp 50	aac Asn	gac Asp	gag Glu	aag Lys	atg Met 55	cgc Arg	cgc Arg	ttc Phe	cag Gln	atc Ile 60	ttc Phe	aag Lys	aac Asn	aac Asn	192
gtg Val 65	aac Asn	cac His	atc Ile	gag Glu	acc Thr 70	ttc Phe	aac Asn	tcc Ser	cgc Arg	aac Asn 75	gag Glu	aac Asn	tcc Ser	tac Tyr	acc Thr 80	240
ctc Leu	ggc Gly	atc Ile	aac Asn	cag Gln 85	ttc Phe	acc Thr	gac Asp	atg Met	acc Thr 90	aac Asn	aac Asn	gag Glu	ttc Phe	atc Ile 95	gcc Ala	288
cag Gln	tac Tyr	acc Thr	ggc Gly 100	ggc Gly	atc Ile	tcc Ser	cgc Arg	ccg Pro 105	ctc Leu	aac Asn	atc Ile	gag Glu	cgc Arg 110	gag Glu	ccg Pro	336
gtg Val	gtg Val	tcc Ser 115	ttc Phe	gac Asp	gac Asp	gtg Val	gac Asp 120	atc Ile	tcc Ser	gcc Ala	gtg Val	ccg Pro 125	cag Gln	tcc Ser	atc Ile	384
gac Asp	tgg Trp 130	cgc Arg	gac Asp	tac Tyr	ggc Gly	gcc Ala 135	gtg Val	acc Thr	tcc Ser	gtg Val	aag Lys 140	aac Asn	cag Gln	aac Asn	ccg Pro	432
tgc Cys 145	ggc Gly	gcc Ala	tgc Cys	tgg Trp	gcc Ala 150	ttc Phe	gcc Ala	gcc Ala	atc Ile	gcc Ala 155	acc Thr	gtg Val	gag Glu	tcc Ser	atc Ile 160	480
tac Tyr	aag Lys	atc Ile	aag Lys	aag Lys 165	ggc Gly	atc Ile	ctc Leu	gag Glu	ccg Pro 170	ctc Leu	tcc Ser	gag Glu	cag Gln	cag Gln 175	gtg Val	528
ctc	gac	tgc	gcc	aag	ggc	tac	ggc	tgc	aag	ggc	ggc	tgg	gag	ttc	cgc	576

Leu Asp Cys	Ala Lys G	y Tyr Gl	/ Cys 185	Lys Gly	Gly 1	Trp Glu 190	Phe	Arg
gcc ttc gag Ala Phe Glu 195	ttc atc at Phe Ile I	c tcc aa e Ser As: 20	ı Lys	ggc gto Gly Val	. Ala S	tcc ggc Ser Gly 205	gcc Ala	atc 624 Ile
tac ccg tac Tyr Pro Tyr 210	aag gcc gc Lys Ala A	cc aag gg a Lys Gl 215	acc Thr	tgc aaq Cys Lys	acc of Thr 7	gac ggc Asp Gly	gtg Val	ccg 672 Pro
aac tee gee Asn Ser Ala 225	tac atc ac Tyr Ile Ti	or Gly Ty	gcc Ala	cgc gtc Arg Val 239	Pro A	cgc aac Arg Asn	aac Asn	gag 720 Glu 240
tcc tcc atg Ser Ser Met	atg tac go Met Tyr A. 245	cc gtg tc la Val Se	aag Lys	cag cco Gln Pro 250	g atc a	acc gtg Thr Val	gcc Ala 255	gtg 768 Val
gac gen aac Asp Ala Asn	gec aac t Ala Asn Pl 260	c cag ta ne Gln Ty	tac r Tyr 265	aag too Lys Se	ggc o	gtg ttc Val Phe 270	aac Asn	ggc 816 Gly
ccg tgc ggc Pro Cys Gly 275	acc tcc c Thr Ser Le	c aac ca eu Asn Hi 28	s Ala	gtg acc Val Th:	Ala	atc ggc Ile Gly 285	tac Tyr	ggc 864 Gly
cag gac tcc Gln Asp Ser 290	atc atc to	ac ccg aa yr Pro Ly 295	g aag s Lys	tgg ggo Trp Gl	gcc a Ala 300	aag tgg Lys Trp	ggc Gly	gag 912 Glu
gcc ggc tac Ala Gly Tyr 305	Ile Arg M	g gcc cg et Ala Ar 10	c gac g Asp	gtg tcc Val Se: 31	ser:	tcc tcc Ser Ser	ggc Gly	atc 960 Ile 320
tgc ggc atc Cys Gly Ile	gcc atc g Ala Ile A 325	ac ccg ct sp Pro Le	c tac u Tyr	ccg ace Pro Th 330	ctc (Leu (gag gag Glu Glu	tag	1005
<210> 70 <211> 334 <212> PRT <213> Arti	ficial Seq	Jence						
<220> <223> Synt	hetic Cons	ruct						
<400> 70								
Met Ala Trp 1	Lys Val G 5	ln Val Va	l Phe	Leu Pho	e Leu	Phe Leu	Cys 15	Val
Met Trp Ala	Ser Pro S	er Ala Al	a Ser 25	Ala As _i	o Glu	Pro Ser 30	Asp	Pro

Page 70

Met Met Lys Arg Phe Glu Glu Trp Met Val Glu Tyr Gly Arg Val Tyr 35 40 45

Lys Asp Asn Asp Glu Lys Met Arg Arg Phe Gln Ile Phe Lys Asn Asn 50 60

Val Asn His Ile Glu Thr Phe Asn Ser Arg Asn Glu Asn Ser Tyr Thr 65 70 75 80

Leu Gly Ile Asn Gln Phe Thr Asp Met Thr Asn Asn Glu Phe Ile Ala 85 90 95

Gin Tyr Thr Gly Gly Ile Ser Arg Pro Leu Asn Ile Glu Arg Glu Pro 100 105 110

Val Val Ser Phe Asp Asp Val Asp Ile Ser Ala Val Pro Gln Ser Ile 115 120 125

Asp Trp Arg Asp Tyr Gly Ala Val Thr Ser Val Lys Asn Gln Asn Pro 130 135 140

Cys Gly Ala Cys Trp Ala Phe Ala Ala Ile Ala Thr Val Glu Ser Ile 145 150 155 160

Tyr Lys Ile Lys Lys Gly Ile Leu Glu Pro Leu Ser Glu Gln Gln Val 165 170 175

Leu Asp Cys Ala Lys Gly Tyr Gly Cys Lys Gly Gly Trp Glu Phe Arg 180 185 190

Ala Phe Glu Phe Ile Ile Ser Asn Lys Gly Val Ala Ser Gly Ala Ile 195 200 205

Tyr Pro Tyr Lys Ala Ala Lys Gly Thr Cys Lys Thr Asp Gly Val Pro 210 220

Asn Ser Ala Tyr Ile Thr Gly Tyr Ala Arg Val Pro Arg Asn Asn Glu 225 230 235 240

Ser Ser Met Met Tyr Ala Val Ser Lys Gln Pro Ile Thr Val Ala Val 245 250 255

Asp Ala Asn Ala Asn Phe Gln Tyr Tyr Lys Ser Gly Val Phe Asn Gly 260 265 270

Page 71

```
Pro Cys Gly Thr Ser Leu Asn His Ala Val Thr Ala Ile Gly Tyr Gly
Gin Asp Ser Ile Ile Tyr Pro Lys Lys Trp Gly Ala Lys Trp Gly Glu
                          295
Ala Gly Tyr Ile Arg Met Ala Arg Asp Val Ser Ser Ser Gly Ile
                                       315
                     310
Cys Gly Ile Ala Ile Asp Pro Leu Tyr Pro Thr Leu Glu Glu
                                       330
                 325
<210> 71
<211> 78
<212> DNA
<213> Artificial Sequence
<220>
<223> Bromealin signal sequence
atggcctgga aggtgcaggt ggtgttcctc ttcctcttcc tctgcgtgat gtgggcctcc
                                                                            60
                                                                            78
cogtocgccg cotcogcc
<210> 72
<211> 26
<212> PRT
<213> Artificial Sequence
<223> Bromealin signal peptide
<400> 72
Met Ala Trp Lys Val Gln Val Val Phe Leu Phe Leu Phe Leu Cys Val
Met Trp Ala Ser Pro Ser Ala Ala Ser Ala
         20
<210> 73
<211> 1050
<212> DNA
<213> Artificial Sequence
<220>
<223> pSYN11000
```

<400> 73 atggcctgga	aggtgcaggt	ggtgttcctc	ttcctcttcc	tctgcgtgat	gtgggcctcc	60
ccgtccgccg	cctccgcgga	cgagccgtcc	gacccgatga	tgaagcgctt	cgaggagtgg	120
atggtggagt	acggccgcgt	gtacaaggac	aacgacgaga	agatgcgccg	cttccagatc	180
ttcaagaaca	acgtgaacca	catcgagacc	ttcaactccc	gcaacgagaa	ctcctacacc	240
ctcggcatca	accagttcac	cgacatgacc	aacaacgagt	tcatcgccca	gtacaccggc	. 300
ggcatctccc	gcccgctcaa	catcgagcgc	gagccggtgg	tgtccttcga	cgacgtggac	360
atctccgccg	tgccgcagtc	catcgactgg	cgcgactacg	gcgccgtgac	ctccgtgaag	420
aaccagaacc	catacaacac	ctgctgggcc	ttcgccgcca	tcgccaccgt	ggagtccatc	480
tacaagatca	agaagggcat	cctcgagccg	ctctccgagc	agcaggtgct	cgactgcgcc	540
aagggctacg	gctgcaaggg	cggctgggag	ttccgcgcct	tcgagttcat	catctccaac	600
aagggcgtgg	cctccggcgc	catctacccg	tacaaggccg	ccaagggcac	ctgcaagacc	660
gacggcgtgc	cgaactccgc	ctacatcacc	ggctacgccc	gcgtgccgcg	caacaacgag	720
tcctccatga	tgtacgccgt	gtccaagcag	ccgatcaccg	tggccgtgga	cgccaacgcc	780
aacttccagt	actacaagtc	cggcgtgttc	aacggcccgt	gcggcacctc	cctcaaccac	840
gccgtgaccg	ccatcggcta	cggccaggac	tccatcatct	acccgaagaa	gtggggcgcc	900
aagtggggcg	aggccggcta	catecgeatg	gcccgcgacg	tgtcctcctc	ctccggcatc	960
tgcggcatcg	ccatcgaccc	getetacceg	accctcgagg	aggtgttcgc	cgaggccatc	1020
gccgccaact	ccaccctcgt	ggccgagtag				1050
<220>	7 ificial Sequ N11589	ience				
<400> 74 tggcctggaa	ggtgcaggtg	gtgttcctct	tcctcttcct	ctgcgtgatg	tgggcctccc	60
cgtccgccgc	ctccgcctcc	tcctcctcct	tcgccgactc	caacccgatc	cgcccggtga	120
ccgaccgcgc	cgcctccacc	gacgagccgt	ccgacccgat	gatgaagcgc	ttcgaggagt	180
ggatggtgga	gtacggccgc	gtgtacaagg	acaacgacga	gaagatgcgc	cgcttccaga	240
tcttcaagaa	caacgtgaac	cacatcgaga	ccttcaactc	ccgcaacgag	aactcctaca	300
ccctcggcat	caaccagttc	accgacatga	ccaacaacga	gttcatcgcc	cagtacaccg	360

gcggcatete ecgecegete aacategage gegageege	gt ggtgtccttc gacgacgtgg 420
acateteege egtgeegeag tecategaet ggegegaet	ca eggegeegtg aceteegtga 480
agaaccagaa cccgtgcggc gcctgctggg ccttcgccç	go catogodaco gtggagtoda 540
tctacaagat caagaagggc atcctcgagc cgctctccg	ga gcagcaggtg ctcgactgcg 600
ccaagggcta cggctgcaag ggcggctggg agttccgcg	go ottogagito atoatotoca 660
acaagggcgt ggcctccggc gccatctacc cgtacaagg	go ogocaagggo acotgoaaga 720
ccgacggcgt gccgaactcc gcctacatca ccggctacg	ge eegegtgeeg egeaacaaeg 780
agtectedat gatgtacged gtgtccaage ageogatea	ac cgtggccgtg gacgccaacg 840
ccaacttcca gtactacaag tccggcgtgt tcaacggc	cc gtgcggcacc tccctcaacc 900
acgoogtgac ogcoatoggo tacggocagg actocatos	at ctacccgaag aagtggggcg 960
ccaagtgggg cgaggccggc tacatccgca tggcccgcg	ga cgtgtcctcc tcctccggca 1020
totgoggoat ogcoatogae cogototace egacceteg	ga ggagtag 1067
<210> 75 <211> 1023	
<212> DNA <213> Artificial Sequence <220> <223> pSYN11587 Sequence	
<212> DNA <213> Artificial Sequence <220>	cc totgogtgat gtgggootoc 60
<212> DNA <213> Artificial Sequence <220> <223> pSYN11587 Sequence <400> 75	3.333
<212> DNA <213> Artificial Sequence <220> <223> pSYN11587 Sequence <400> 75 atggcctgga aggtgcaggt ggtgttcctc ttcctcttc	ga tgaagcgctt cgaggagtgg 120
<212> DNA <213> Artificial Sequence <220> <223> pSYN11587 Sequence <400> 75 atggcctgga aggtgcaggt ggtgttcctc ttcctcttc ccgtccgccg cctccgcgga cgagccgtcc gacccgatg	ga tgaagcgctt cgaggagtgg 120 ga agatgcgccg cttccagatc 180
<pre><212> DNA <213> Artificial Sequence <220> <223> pSYN11587 Sequence <400> 75 atggcctqqa aggtgcaggt ggtgttcctc ttcctcttc ccqtccqccq cctccqcqqa cgaqccqtcc gacccqatq atgqtqqagt acgqccqcqt gtacaaggac aacgacqaq</pre>	ga tgaagcgctt cgaggagtgg 120 ga agatgcgccg cttccagatc 180 cc gcaacgagaa ctcctacacc 240
<pre><212> DNA <213> Artificial Sequence <220> <223> pSYN11587 Sequence <400> 75 atggcctqqa aggtgcaggt ggtgttcctc ttcctcttc ccgtccgccg cctccgcgga cgagccgtcc gacccgatc atggtggagt acggccgcgt gtacaaggac aacgacgac ttcaagaaca acgtgaacca catcgagacc ttcaactcc</pre>	ga tgaagcgctt cgaggagtgg 120 ga agatgcgccg cttccagatc 180 cc gcaacgagaa ctcctacacc 240 gt tcatcgccca gtacaccggc 300
<pre><212> DNA <213> Artificial Sequence <220> <223> pSYN11587 Sequence <400> 75 atggcctqqa aggtgcaggt ggtgttcctc ttcctcttc ccgtccgccg cctccgcgga cgagccgtcc gacccgatc atggtggagt acggccgct gtacaaggac aacgacgac ttcaagaaca acgtgaacca catcgagacc ttcaactcc ctcggcatca accagttcac cgacatgacc aacaacgag</pre>	ga tgaagcgctt cgaggagtgg 120 ga agatgcgccg cttccagatc 180 gc gcaacgagaa ctcctacacc 240 gt tcatcgccca gtacaccggc 300 gg tgtccttcga cgacgtggac 360
<pre><212> DNA <213> Artificial Sequence <220> <223> pSYNl1587 Sequence <400> 75 atggcctgga aggtgcaggt ggtgttcctc ttcctcttc ccgtccgccg cctccgcgga cgagccgtcc gacccgatg atggtggagt acggccgcgt gtacaaggac aacgacggg ttcaagaaca acgtgaacca catcgagacc ttcaactcc ctcggcatca accagttcac cgacatgacc aacaacgag ggcatctccc gcccgctcaa catcgagcgc gagccggtg</pre>	ga tgaagcgctt cgaggagtgg 120 ga agatgegeeg ettecagate 180 gc gcaacgagaa etcetacace 240 gt teategeeca gtacacegge 300 gg tgteettega egaegtggae 360 gg gegeegtgae etcegtgaag 420
<pre><212> DNA <213> Artificial Sequence <220> <223> pSYN11587 Sequence <400> 75 atggcctgga aggtgcaggt ggtgttcctc ttcctcttc ccgtccgccg cctccgcgga cgagccgtcc gacccgatg atggtggagt acggccgcgt gtacaaggac aacgacgag ttcaagaaca acgtgaacca catcgagacc ttcaactcc ctcggcatca accagttcac cgacatgacc aacaacgag ggcatctccc gcccgctcaa catcgagcgc gagccggtg atctccgccg tgccgcagtc catcgactgg cgcgactac</pre>	ga tgaagcgctt cgaggagtgg 120 ga agatgcgccg cttccagatc 180 gc gcaacgagaa ctcctacacc 240 gt tcatcgccca gtacaccggc 300 gg tgtccttcga cgacgtggac 360 gg gcgccgtgac ctccgtgaag 420 ga tcgccaccgt ggagtccatc 480
<pre><212> DNA <213> Artificial Sequence <220> <223> pSYN11587 Sequence <400> 75 atggcctgga aggtgcaggt ggtgttcctc ttcctcttc ccgtccgccg cctccgcgga cgagccgtcc gacccgatg atggtggagt acggccgct gtacaaggac aacgacgag ttcaagaaca acgtgaacca catcgagacc ttcaactcc ctcggcatca accagttcac cgacatgacc aacaacgag ggcatctccc gcccgctcaa catcgagcgc gagccggtg atctccgccg tgccgcagtc catcgactgg cgcgactac aaccagaacc cgtgcgcgc ctgctgggcc ttcgccgcc</pre>	ga tgaagcgctt cgaggagtgg 120 ga agatgcgccg cttccagatc 180 gc gcaacgagaa ctcctacacc 240 gt tcatcgccca gtacaccggc 300 gg tgtccttcga cgacgtggac 360 gg gcgccgtgac ctccgtgaag 420 ga tcgccaccgt ggagtccatc 480 gc agcaggtgct cgactgcgcc 540
<pre><212> DNA <213> Artificial Sequence <220> <223> pSYN11587 Sequence <400> 75 atggcctqqa aggtgcaggt ggtgttcctc ttcctcttc ccqtccqccq cctccqcqqa cqaqccqtcc gacccqatq atgqtqqaqt acgqccqcqt gtacaaqqac aacqacqaq ttcaaqaaca acqtqaacca catcqaqacc ttcaactcc ctcqqcatca accaqttcac cqacatqacc aacaacqaq gqcatctccc qcccqctcaa catcqaqcqc qaqccqqtq atctccqccq tgccqcaqtc catcqactqq cqcqactac aaccaqaacc cqtqcqqcqc ctqctqqqcc ttcqccqcc tacaaqatca agaaqqqcat cctcqaqccq ctctccqaq</pre>	ga tgaagcgctt cgaggagtgg 120 ga agatgcgccg cttccagatc 180 gc gcaacgagaa ctcctacacc 240 gt tcatcgccca gtacaccggc 300 gg tgtccttcga cgacgtggac 360 gg gcgccgtgac ctccgtgaag 420 ga tcgccaccgt ggagtccatc 480 gc agcaggtgct cgactgcgcc 540 gt tcgagttcat catctccaac 600

tcctcca:	tga	tgtacgccgt	gtccaagcag	ccgatcaccg	tggccgtgga	cgccaacgcc	780
aacttcca	agt	actacaagtc	cggcgtgttc	aacggcccgt	gcggcacctc	cctcaaccac	840
gccgtga	ccg	ccatoggota	cggccaggac	tccatcatct	acccgaagaa	gtggggcgcc	900
aagtggg	geg	aggccggcta	cateegeatg	gcccgcgacg	tgtcctcctc	ctccggcatc	960
tgcggca	tcg	ccategacee	gctctacccg	accctcgagg	agtccgagaa	ggacgagctg	1020
tag							1023
<211> <212>	76 990 DNA Arti	ficial Sequ	uence				
<223>	pSYN	112169 Seque	ence				
<400> atgaggg	76 tgt	tgctcgttgc	cctcgctctc	ctggctctcg	ctgcgagcgc	cacctccatg	60
gcggacg	age	egteegacee	gatgatgaag	cgcttcgagg	agtggatggt	ggagtacggc	120
egegt gt	aca	açgacaacga	cgagaagatg	cgccgcttcc	agatetteaa	gaacaacgtg	180
aaccaca	tcg	agaccttcaa	ctcccgcaac	gagaactcct	acaccctcgg	catcaaccag	240
ttcaccg	aca	tgaccaacaa	cgagttcatc	gcccagtaca	ccggcggcat	ctcccgcccg	300
ctcaaca	tcg	agcgcgagcc	ggtggtgtcc	ttcgacgacg	tggacatctc	cgccgtgccg	360
cagtoca	tcg	actggcgcga	ctacggcgcc	gtgacctccg	tgaagaacca	gaacccgtgc	420
ggcgcct	gct	gggccttcgc	cgccatcgcc	accgtggagt	ccatctacaa	gatcaagaag	480
ggcatcc	tcģ	agccgctctc	cgagcagcag	gtgctcgact	gcgccaaggg	ctacggctgc	540
aagggcg	ąct	gggagttccg	cgccttcgag	ttcatcatct	ccaacaaggg	cgtggcctcc	600
ggcgcca	tct	acccgtacaa	ggccgccaag	ggcacctgca	agaccgacgg	cgtgccgaac	660
teegeet	аса	tcaccggcta	cgcccgcgtg	ccgcgcaaca	acgagtcctc	catgatgtac	720
geegtgt	cca	agcagccgat	caccgtggcc	gtggacgcca	acgccaactt	ccagtactac	780
					accacgccgt		840
ggetacg	gcc	aggactccat	catctacccg	aagaagtggg	gcgccaagtg	gggcgaggcc	900
ggctaca	tcc	gcatggcccg	cgacgtgtcc	tcctcctccg	gcatctgcgg	catcgccatc	960
gacccgc	tct	accegaccet	cgaggagtag				990

```
<210> 77
<211> 1170
<212> DNA
<213> Artificial Sequence
<220>
<223> pSYN12575 Sequence
atgctggcgg ctctggccac gtcgcagctc gtcgcaacgc gcgccggcct gggcgtcccg
                                                                    . 60
gacqcgtcca cgttccgccg cggcgccgcg cagggcctga ggggggcccg ggcgtcggcg
                                                                     120
gcggcggaca cgctcagcat gcggaccagc gcgcgcgcgg cgcccaggca ccagcaccag
                                                                     180
caggogogoc goggagocag gitocogica elegiogist gogocagogo oggogocata
                                                                     240
geggaegage egteegaeee gatgatgaag egettegagg agtggatggt ggagtaegge
                                                                     300
                                                                     360
egegtgtaca aggacaacga egagaagatg egeegettee agatetteaa gaacaacgtg
aaccacateg agacetteaa eteeegcaae gagaaeteet acaceetegg cateaaccag
                                                                     420
                                                                     480
ttcaccgaca tgaccaacaa cgagttcatc gcccagtaca ccggcggcat ctcccgcccg
ctcaacatcg agogogaged ggtggtgted ttcgacgacg tggacatete egocgtgccg
                                                                     540
cagtocateg actggegega ctaeggegee gtgaceteeg tgaagaacea gaaceegtge
                                                                     600
ggcgcctgct gggccttcgc cgccatcgcc accgtggagt ccatctacaa gatcaagaag
                                                                     660
ggcatecteg ageogetete egageageag gtgetegaet gegeeaaggg etaeggetge
                                                                     720
aagggcggct gggagttccg cgccttcgag ttcatcatct ccaacaaggg cgtggcctcc
                                                                     780
ggogocatot accogtacaa ggoogocaag ggoacotgoa agacogaogg ogtgoogaac
                                                                     840
toogeotaca teacoggera egeoegogtg cogogoaaca acgagtecte catgatgtac
                                                                     900
                                                                     960
gccqtgtcca agcagccgat caccgtggcc gtggacgcca acgccaactt ccagtactac
aagteeggeg tgttcaaegg eeegtgegge aceteeetea accaegeegt gacegeeate
                                                                    1020
                                                                    1080
ggotacggod aggactodat catotaccog aagaagtggg gogodaagtg gggogaggod
ggotacatoo gcatggcccg cgacgtgtoo tootootoog gcatotgogg catcgccato
                                                                    1140
                                                                    1170
gacccoctct acccgaccct cgaggagtag
<210> 78
```

<210> 78 <211> 1068

<212> DNA <213> Artificial Sequence

<220>

<223> pSM270 Sequence

.100. 30	
<400> 78 atggcctgga aggtgcaggt ggtgtteete tteetettee tetgegtgat gtgggcetee	60
cogrecged ecreeged etected tregogact ccaaccegat ccgcccggtg	120
accgaccgcg ccgcctccac cgacgagecg tecgacccga tgatgaageg cttegaggag	180
tggatggtgg agtacggccg cgtgtacaag gacaacgacg agaagatgcg ccgcttccag	240
atottoaaga acaacgtgaa coacatogag acottoaact coogcaacga gaactootac	300
acceteggea teaaceagtt cacegacatg accaacaacg agtteatege ecagtacace	360
ggcggcatct cccgcccgct caacatcgag cgcgagccgg tggtgtcctt cgacgacgtg	420
gacateteeg cegtgeegea greeategae tggegegaet aeggegeegt gaceteegtg	480
aagaaccaga accegtgegg egeetgetgg geettegeeg ceategeeac egtggagtee	540
atotacaaga toaagaaggg catootogag cogototoog agcagcaggt gotogactgo	600
gccaagggct acggctgcaa gggcggctgg gagttccgcg cettcgagtt catcatetec	660
aacaagggcg tggcctccgg cgccatctac ccgtacaagg ccgccaaggg cacctgcaag	720
accgaeggeg tgeegaacte egeevacate aceggetacg ecegegtgee gegeaacaac	780
gagtcctcca tgatgtacgc cgtgtccaag cagccgatca ccgtggccgt ggacgccaac	840
gccaacttcc agtactacaa gtccggcgtg ttcaacggcc cgtgcggcac ctccctcaac	900
cacqccqtga ccqccatcqq ctacqqccaq gactccatca tctacccqaa gaagtqqqqc	960
godaagtggg gogaggoogg ctacatoogo atggooogog acgtgtooto otootooggo	1020
atotgoogoa togooatoga coogototac oogacootog aggagtag	1068
<210> 79 <211> 1497 <212> DNA <213> Trichoderma reesei	
<220> <221> CDS <222> (1)(1497) <223> Trichoderma reesei cellobiohyrodlase I	
<400> 79 atg cag tog gcg tgt act ctc caa tog gag act cac ccg cct ctg aca Met Gln Ser Ala Cys Thr Leu Gln Ser Glu Thr His Pro Pro Leu Thr 1 5 10 15	48
tgg cag aaa tgc tcg tct ggt ggc acg tgc act caa cag aca ggc tcc Trp Gln Lys Cys Ser Ser Gly Gly Thr Cys Thr Gln Gln Thr Gly Ser 20 25 30	96

gtg Val	gtc Val	atc Ile 35	gac Asp	gcc Ala	aac Asn	tgg Trp	cgc Arg 40	tgg Trp	act Thr	cac His	gct Ala	acg Thr 45	aac Asn	agc Ser	agc Ser	144
acg Thr	aac Asn 50	t gc Cys	tac Tyr	gat Asp	G] À	aac Asn 55	act Thr	tgg Trp	agc Ser	tcg Ser	acc Thr 60	cta Leu	tgt Cys	cct Pro	gac Asp	192
aac Asn 65	gag Glu	acc Thr	tgc Cys	gcg Ala	aag Lys 70	aac Asn	t gc Cys	tgt Cys	ctg Leu	gac Asp 75	ggt Gly	gcc Ala	gcc Ala	tac Tyr	gcg Ala 80	240
t cc Ser	acg Thr	tac Tyr	ggā Gly	gtt Val 85	acc Thr	acg Thr	agc Ser	ggt Gly	aac Asn 90	agc Ser	ctc Leu	tcc Ser	att Ile	ggc Gly 95	ttt Phe	288
gtc Val	acc Thr	cag Gln	tct Ser 100	gcg Ala	cag Gln	aag Lys	aac Asn	gtt Val 105	ggc Gly	gct Ala	cgc Arg	ctt Leu	tac Tyr 110	ctt Leu	atg Met	336
gcg Ala	agc Ser	gac Asp 115	acg Thr	acc Thr	tac Tyr	cag Gln	gaa Glu 120	ttc Phe	acc Thr	ctg Leu	ctt Leu	ggc Gly 125	aac Asn	gag Glu	ttc Phe	384
tct Ser	ttc Phe 130	gat Asp	gtt Val	gat Asp	gtt Val	tcg Ser 135	cag Gln	ctg Leu	ccg Pro	tgc Cys	ggc Gly 140	t t g Leu	aac Asn	gga Gly	gct Ala	432
ctc Leu 145	tac Tyr	ttc Phe	gtg Val	tcc Ser	atg Met 150	gac Asp	gcg Ala	gat Asp	ggt Gly	ggc Gly 155	gtg Val	agc Ser	aag Lys	tat Tyr	ccc Pro 160	480
acc Thr	aac Asn	acc Thr	gct Ala	ggc Gly 165	gcc Ala	aag Lys	tac Tyr	ggc Gly	acg Thr 170	ggg Gly	tac Tyr	tgt Cys	gac Asp	agc Ser 175	cag Gln	528
tgt Cys	ccc Pro	cgc Arg	gat Asp 180	ctg Leu	aag Lys	ttc Phe	atc Ile	aat Asn 185	ggc Gly	cag Gln	gcc Ala	aac Asn	gtt Val 190	gag Glu	ggc Gly	576
tgg Trp	çag Glu	ccg Pro 195	tca Ser	tcc Ser	aac Asn	aac Asn	gcg Ala 200	aac Asn	acg Thr	ggc Gly	att Ile	gga Gly 205	gga Gly	cac His	gga Gly	624
agc Ser	tgc Cys 210	tgc Cys	tct Ser	gag Glu	atg Met	gat Asp 215	atc Ile	tgg Trp	gag Glu	gcc Ala	aac Asn 220	tcc Ser	atc Ile	tcc Ser	gag Glu	672
gct Ala 225	ct t Leu	acc Thr	ccc Pro	cac His	cct Pro 230	tgc Cys	acg Thr	act Thr	gtc Val	ggc Gly 235	cag Gln	gag Glu	atc Ile	t gc Cys	gag Glu 240	720
ggt Gly	gat Asp	999 Gly	tgc Cys	ggc Gly 245	gga Gly	act Thr	tac Tyr	t c c Se r	gat Asp 250	aac Asn	aga Arg	tat Tyr	ggc Gly	ggc Gly 255	act Thr	768
tgc Cys	gat Asp	ccc Pro	gat Asp	ggc	t gc Cys	gac Asp	tgg Trp	aac Asn	cca Pro	tac Tyr	cgc Arg	ct g Leu	ggc Gly	aac Asn	acc Thr	816

Page 78

	260	265		270	
age tte tac Ser Phe Tyr 275	Gly Pro Gly	tct agc ttt Ser Ser Phe 280	acc ctc gat ac Thr Leu Asp Th 28	r Thr Lys Ly	a 864 s
ttg acc gtt Leu Thr Val 290	gtc acc cag Val Thr Gln	ttc gag acg Phe Glu Thr 295	g tog ggt gcc at Ser Gly Ala Il 300	c aac cga ta e Asn Arg Ty	c 912 r
tat gtc cag Tyr Val Gln 305	aat ggc gtc Asn Gly Val 310	act ttc cag Thr Phe Gln	g cag ccc aac go n Gln Pro Asn Al 315	c gag ctt gg a Glu Leu Gl 32	У
agt tac tot Ser Tyr Ser	ggc aac gag Gly Asn Glu 325	ctc aac gat Leu Asn Asp	gat tac tgc ac Asp Tyr Cys Th 330	a gct gag ga r Ala Glu Gl 335	g 1008 u
gca gaa tto Ala Glu Phe	ggc gga tec Gly Gly Ser 340	tet tte tea Ser Phe Ser 345	a gac aag ggc gg Asp Lys Gly Gl	c ctg act ca y Leu Thr Gl 350	g 1056 n
ttc aag aag Phe Lys Lys 355	Ala Thr Ser	ggc ggc atg Gly Gly Met 360	g gtt ctg gtc at Val Leu Val Me 36	et Ser Leu Tr	g 1104 P
gat gat tac Asp Asp Tyr 370	tac gcc aac Tyr Ala Asn	atg ctg tgg Met Leu Trp 375	g ctg gac tcc ac b Leu Asp Ser Tr 380	cc tac ccg ac or Tyr Pro Th	a 1152 r
aac gag acc Asn Glu Thr 385	tcc tcc aca Ser Ser Thr 390	ccc ggt gcc Pro Gly Ala	gtg cgc gga ag Nal Arg Gly Se 395	gc tgc tcc ac er Cys Ser Th 40	r
age tee ggt Ser Ser Gly	gtc cct gct Val Pro Ala 405	cag gtc gaa Gln Val Glu	a tot cag tot co 1 Ser Gln Ser Pr 410	cc aac gcc aa o Asn Ala Ly 415	g 1248 s
gtc acc ttc Val Thr Phe	tcc aac atc Ser Asn Ile 420	aag ttc gga Lys Phe Gly 425	a ccc att ggc ac / Pro Ile Gly Se 5	gc acc ggc aa er Thr Gly As 430	c 1296 n
cct agc ggc Pro Ser Gly 435	Gly Asn Pro	ccc ggc gga Pro Gly Gly 440	a aac ccg cct go / Asn Pro Pro Gl 44	ly Thr Thr Th	1344 r
acc ege ege Thr Arg Arg 450	cca gcc act Pro Ala Thr	acc act gga Thr Thr Gly 455	a ago tot coo go / Ser Ser Pro Gl 460	ga cct acc ca y Pro Thr Gl	g 1392 n
tot cac tac Ser His Tyr 465	ggc cag tgc Gly Gln Cys 470	ggc ggt att Gly Gly Ile	ggc tac agc gg Gly Tyr Ser Gl 475	gc ccc acg gt y Pro Thr Va 48	1
tge gee age Cys Ala Ser	ggc aca act Gly Thr Thr 485	tgc cag gtc Cys Gln Val	c ctg aac cct ta Leu Asn Pro Ty 490	oc tac tot ca vr Tyr Ser Gl 495	g 1488 n

1497 tgc ctg taa Cys Leu

<210> 80

<211> 498 <212> PRT <213> Trichoderma reesei

<400> 80

Met Gln Ser Ala Cys Thr Leu Gln Ser Glu Thr His Pro Pro Leu Thr

Trp Gln Lys Cys Ser Ser Gly Gly Thr Cys Thr Gln Gln Thr Gly Ser

Val Val Ile Asp Ala Asn Trp Arg Trp Thr His Ala Thr Asn Ser Ser

Thr Asn Cys Tyr Asp Gly Asn Thr Trp Ser Ser Thr Leu Cys Pro Asp

Asn Glu Thr Cys Ala Lys Asn Cys Cys Leu Asp Gly Ala Ala Tyr Ala

Ser Thr Tyr Gly Val Thr Thr Ser Gly Asn Ser Leu Ser Ile Gly Phe

Val Thr Gln Ser Ala Gln Lys Asn Val Gly Ala Arg Leu Tyr Leu Met 105

Ala Ser Asp Thr Thr Tyr Gln Glu Phe Thr Leu Leu Gly Asn Glu Phe

Ser Phe Asp Val Asp Val Ser Gln Leu Pro Cys Gly Leu Asn Gly Ala

Leu Tyr Phe Val Ser Met Asp Ala Asp Gly Gly Val Ser Lys Tyr Pro

Thr Asn Thr Ala Gly Ala Lys Tyr Gly Thr Gly Tyr Cys Asp Ser Gln

Cys Pro Arg Asp Leu Lys Phe Ile Asn Gly Gln Ala Asn Val Glu Gly 185

Page 80

rrp	Glu	Pro 195	Ser	Ser	Asn	Asn	Ala 200	Asn	Thr	Gly	Ile	Gly 205	Gly	His	Gly
Ser	Cys 210	Cys	Ser	Glu	Met	Asp 215	Ile	Trp	Glu	Ala	Asn 220	Ser	Ile	Ser	Glu
Ala 225	Leu	Thr	Pro	His	Pro 230	Cys	Thr	Thr	Val	Gly 235	Gln	Glu	Ile	Cys	Glu 240
Gly	Asp	Gly	Cys	Gly 245	Gly	Thr	Туr	Ser	Asp 250	Asn	Arg	Tyr	Gly	Gly 255	Thr
Cys	Asp	Pro	Asp 260	Gly	Cys	Asp	Trp	Asn 265	Pro	Tyr	Arg	Leu	Gly 270	Asn	Thr
Ser	Phe	Tyr 275	Gly	Pro	Gly	Ser	Ser 280	Phe	Thr	Leu	Asp	Thr 285	Thr	Lys	Lys
Leu	Thr 290	Val	Val	Thr	Gln	Phe 295	Glu	Thr	Ser	Gly	Ala 300	Ile	Asn	Arg	Tyr
Tyr 305	Val	Gln	Asn	Gly	Val 310	Thr	Phe	Gln	Gln	Pro 315	Asn	Ala	Glu	Leu	Gly 320
Ser	Tyr	Ser	Gly	Asn 325	Glu	Leu	Asn	Asp	Asp 330	Tyr	Cys	Thr	Ala	Glu 335	Glu
Ala	Glu	Phe	Gly 340	Gly	Ser	Ser	Phe	Ser 345	Asp	Lys	Gly	Gly	Leu 350	Thr	Gln
Phe	Lys	Lys 355	Ala	Thr	Ser	Gly	Gly 360	Met	Val	Leu	Val	Met 365	Ser	Leu	Trp
Asp	Asp 370		Tyr	Ala	Asn	Met 375	Leu	Trp	Leu	Asp	Ser 380	Thr	Tyr	Pro	Thr
Asn 385	Glu	Thr	Ser	Ser	Thr 390	Pro	Gly	Ala	Val	Arg 395	Gly	Ser	Cys	Ser	Thr 400
Ser	Ser	Gly	Val	Pro	Ala	Gln	Val	Glu	Ser 410	Gln	Ser	Pro	Asn	Ala 415	Lys

420	Asn Ile	Lys Phe	Gly Pro 425	Ile Gly	Ser	Thr 430	Gly	Asn	
Pro Ser Gly Gly 435	/ Asn Pro	Pro Gly 440	Gly Asn	Pro Pro	Gly 445	Thr	Thr	Thr	
Thr Arg Arg Pro	o Ala Thr	Thr Thr 455	Gly Ser	Ser Pro 460	Gly	Pro	Thr	Gln .	
Ser His Tyr Gly 465	y Gln Cys 470	Gly Gly	Ile Gly	Tyr Ser 475	Gly	Pro	Thr	Val 480	
Cys Ala Ser Gly	y Thr Thr 485	Cys Gln	Val Leu 490	Asn Pro	Tyr	Tyr	Ser 495	Gln	
Cys Leu									
<210> 81 <211> 1365 <212> DNA <213> Trichodo	erma reese	: 1							
<220> <221> CDS <222> (1)(1 <223> trichod	365) erma ræese	ei cello	biohydro	olase II					
<221> CDS <222> (1)(1	erma reese	coo caa	act tag	tca ago	gtc Val	tgg Trp	ggc Gly 15	caa 48 Gìn	
<pre><221> CDS <222> (1)(1 <223> trichod <400> 81 atg gtg cct ct Met Val Pro Le</pre>	erma reese a gag gag u Glu Glu 5 g aat tgg n Asn Trp	cgg caa Arg Gln	gct tgc Ala Cys 10	tca ago	vai ,	tcc	15 gga	Gln agc 96	
<pre><221> CDS <222> (1)(1 <223> trichod <400> 81 atg gtg cct ct Met Val Pro Le ! tgt ggt ggc ca Cys Gly Gly Gl</pre>	erma reese a gag gag u Glu Glu 5 g aat tgg n Asn Trp	cgg caa Arg Gln tcg ggt Ser Gly	gct tgc Ala Cys 10 ccg act Pro Thi 25	tca ago Ser Ser tgc tgt Cys Cys	gct Ala	tcc Ser 30	gga Gly ggc	Gln agc 96 Ser gct 144	
<pre><221> CDS <222> (1)(1 <223> trichod <400> 81 atg gtg cct ct. Met Val Pro Le ! tct ggt ggc ca Cys Gly Gly Gl 20 aca tgc gtc ta Thr Cys Val Ty</pre>	a gag gag u Glu Glu 5 g aat tgg n Asn Trp c tcc aac r Ser Asn	cgg caa Arg Gln tcg ggt Ser Gly gac tat Asp Tyr 40	gct tgc Ala Cys 10 ccg act Pro Thi 25 tac tcc Tyr Sei	tca ago Ser Ser tgc tgt Cys Cys cag tgt Gln Cys	gct Ala ctt Leu 45	tcc Ser 30 ccc Pro	gga Gly ggc Gly	Gln agc 96 Ser gct 144 Ala gta 192	
<pre><221> CDS <222> (1)(1 <223> trichod <400> 81 atg gtg cct ct Met Val Pro Le ! tct ggt ggc ca Cys Gly Gly Gl 20 aca tgc gtc ta Thr Cys Val Ty 35 gca agc tca ag Ala Ser Ser Se</pre>	a gag gag u Glu Glu 5 g aat tgg n Asn Trp c tcc aac r Ser Asn c tcg tcc r Ser Ser	cgg caa Arg Gln tcg ggt Ser Gly gac tat Asp Tyr 40 acg cgc Thr Arg	gct tgc Ala Cys 10 ccg act Pro Thi 25 tac tcc Tyr Sei gcc gcc Ala Ala	tca ago Ser Ser tgc tgt Cys Cys cag tgt Gln Cys tcg acc Ser Thi 60	gct Ala ctt Leu 45 act Thr	tcc Ser 30 ccc Pro tca Ser	gga Gly ggc Gly cga Arg	Gln agc 96 Ser gct 144 Ala gta 192 Val tct 240	

Page 82

				85					90					95		
ggc Gly	aac Asn	cct Pro	ttt Phe 100	gtt Val	999 Gly	gtc Val	act Thr	cct Pro 105	tgg Trp	gcc Ala	aat Asn	gca Ala	tat Tyr 110	tac Tyr	gcc Ala	336
tct Ser	gaa Glu	gtt Val 115	agc Ser	agc Ser	ctc Leu	gct Ala	att Ile 120	cct Pro	agc Ser	ttg Leu	act Thr	gga Gly 125	gcc Ala	atg Met	gcc Ala	384
act Thr	get Ala 130	gca Ala	gca Ala	gct Ala	gtc Val	gca Ala 135	aag Lys	gtt Val	ccc Pro	tct Ser	ttt Phe 140	atg Met	tgg Trp	cta Leu	gat Asp	432
acu Thr 145	ctt Leu	gac Asp	aag Lys	acc Thr	cct Pro 150	ctc Leu	atg Met	gag Glu	caa Gln	acc Thr 155	ttg Leu	gcc Ala	gac Asp	atc Ile	cgc Arg 160	480
acc Thr	gcc Ala	aac Asn	aag Lys	aat Asn 165	ggc Gly	ggt Gly	aac Asn	tat Tyr	gcc Ala 170	gga Gly	cag Gln	ttt Phe	gtg Val	gtg Val 175	tat Tyr	528
gac Asp	t t g Le u	ccg Pro	gat Asp 180	cgc Arg	gat Asp	tgc Cys	gct Ala	gcc Ala 185	ctt Leu	gcc Ala	tcg Ser	aat Asn	ggc Gly 190	gaa Glu	tac Tyr	576
tct Ser	att Ile	gcc Ala 195	gat Asp	ggt Gly	ggc Gly	gtc Val	gcc Ala 200	aaa Lys	tat Tyr	aag Lys	aac Asn	tat Tyr 205	atc Ile	gac Asp	acc Thr	624
att Ile	egt Arg 210	caa Gln	att Ile	gtc Val	gtg Val	gaa Glu 215	tat Tyr	tcc Ser	gat Asp	atc Ile	cgg Arg 220	acc Thr	ctc Leu	ctg Leu	gtt Val	672
att 11e 225	gag Glu	cct Pro	gac Asp	tct Ser	ctt Leu 230	gcc Ala	aac Asn	ctg Leu	gtg Val	acc Thr 235	aac Asn	ctc Leu	ggt Gly	act Thr	cca Pro 240	720
aag Lys	tgt Cys	gcc Ala	aat Asn	gct Ala 245	cag Gln	tca Ser	gcc Ala	tac Tyr	ctt Leu 250	gag Glu	tgc Cys	atc Ile	aac Asn	tac Tyr 255	gcc Ala	768
gtc Val	aca Thr	cag Gln	ctg Leu 260	aac Asn	ctt Leu	cca Pro	aat Asn	gtt Val 265	gcg Ala	atg Met	tat Tyr	ttg Leu	gac Asp 270	gct Ala	ggc Gly	816
cat His	gca Ala	gga Gly 275	tgg Trp	ctt Leu	ggc Gly	tgg Trp	ccg Pro 280	gca Ala	aac Asn	caa Gln	gac Asp	ccg Pro 285	gcc Ala	gct Ala	cag Gln	864
cta Leu	ttt Phe 290	gca Ala	aat Asn	gtt Val	tac Tyr	aag Lys 295	aat Asn	gca Ala	tcg Ser	tct Ser	ccg Pro 300	aga Arg	gct Ala	ctt Leu	cgc Arg	912
gga Gly 3 0 5	t t g Leu	gca Ala	acc Thr	aat Asn	gtc Val 310	gcc Ala	aac Asn	tac Tyr	aac Asn	999 Gly 315	tgg Trp	aac Asn	att Ile	acc Thr	agc Ser 320	960

ccc Pro	cca Pro	t cg Ser	tac Tyr	acg Thr	caa Gln	ggc Gly	aac Asn	gct Ala	Val	tac Tyr	aac Asn	gag Glu	aag Lys	Leu	tac Tyr	1008
			-	325					330					335		
atc Ile	cac His	gct Ala	att Ile 340	gga Gly	cct Pro	ctt Leu	ctt Leu	gcc Ala 345	aat Asn	cac His	ggc Gly	tgg Trp	tcc Ser 350	aac Asn	gcc Ala	1056
ttc Phe	ttc Phe	atc Ile 355	act Thr	gat Asp	caa Gln	ggt Gly	cga Arg 360	tcg Ser	gga Gly	aag Lys	cag Gln	cct Pro 365	acc Thr	gga Gly	cag Gln	1104
caa Gln	cag Gln 370	tgg Trp	gga Gly	gac Asp	tgg Trp	tgc Cys 375	aat Asn	gtg Val	atc Ile	ggc Gly	acc Thr 380	gga Gly	ttt Phe	ggt Gly	att Ile	1152
cgc Arg 385	cca Pro	tcc Ser	gca Ala	aac Asn	act Thr 390	999 Glу	gac Asp	t cg Ser	ttg Leu	ctg Leu 395	gat Asp	tcg Ser	ttt Phe	gtc Val	tgg Trp 400	1200
gte Val	aag Lys	cca Pro	ggc Gly	ggc Gly 405	gag Glu	tgt Cys	gac Asp	ggc Gly	acc Thr 410	agc Ser	gac Asp	agc Ser	agt Ser	gcg Ala 415	cca Pro	1248
cga Arg	ttt Phe	gac Asp	tcc Ser 420	cac His	tgt Cys	gcg Ala	ctc Leu	cca Pro 425	gat Asp	gcc Ala	ttg Leu	caa Gln	ccg Pro 430	gcg Ala	cct Pro	1296
caa Gln	gct Ala	ggt Gly 435	gct Ala	tgg Trp	ttc Phe	caa Gln	gcc Ala 440	tac Tyr	ttt Phe	gtg Val	cag Gln	ctt Leu 445	ctc Leu	aca Thr	aac Asn	1344
gca Ala	aac Asn 450	cca Pro	t cg Ser	ttc Phe	ctg Leu	tag										1365
<21 <21 <21 <21	1 > 2 >	52 454 PRT Tric	hode	rma	rees	ei										
< 40	0 >	82														
Met 1	Val	Pro	Leu	Glu 5	Glu	Arg	Gln	Ala	Cys 10	Ser	Ser	Val	Trp	Gly 15	Gln	
Cys	Gly	Gly	Gln 20	Asn	Trp	Ser	Gly	Pro 25	Thr	Cys	Cys	Ala	Ser 30	Gly	Ser	
Thr	Суѕ	Val 35	Туr	Ser	Asn	Asp	Tyr 40	Tyr	Ser	Gln	Cys	Leu 45	Pro	Gly	Ala	
Ala	Ser 50	Ser	Ser	Ser	Ser	Thr 55	Arg	Ala	Ala	Ser	Thr 60	Thr	Ser	Arg	Val	

Page 84

- Ser Pro Thr Thr Ser Arg Ser Ser Ser Ala Thr Pro Pro Pro Gly Ser 65 70 75 80
- Thr Thr Arg Val Pro Pro Val Gly Ser Gly Thr Ala Thr Tyr Ser 85 90 95
- Gly Asn Pro Phe Val Gly Val Thr Pro Trp Ala Asn Ala Tyr Tyr Ala 100 105 110
- Ser Glu Val Ser Ser Leu Ala Ile Pro Ser Leu Thr Gly Ala Met Ala 115 120 125
- Thr Ala Ala Ala Ala Val Ala Lys Val Pro Ser Phe Met Trp Leu Asp 130 135 140
- Thr Leu Asp Lys Thr Pro Leu Met Glu Gln Thr Leu Ala Asp Ile Arg 145 150 155 160
- Thr Ala Asn Lys Asn Gly Gly Asn Tyr Ala Gly Gln Phe Val Val Tyr 165 170 175
- Asp Leu Pro Asp Arg Asp Cys Ala Ala Leu Ala Ser Asn Gly Glu Tyr 180 185
- Ser Ile Ala Asp Gly Gly Val Ala Lys Tyr Lys Asn Tyr Ile Asp Thr
- Ile Arg Gln Ile Val Val Glu Tyr Ser Asp Ile Arg Thr Leu Leu Val 210 220
- Ile Glu Pro Asp Ser Leu Ala Asn Leu Val Thr Asn Leu Gly Thr Pro 225 230 240
- Lys Cys Ala Asn Ala Gln Ser Ala Tyr Leu Glu Cys Ile Asn Tyr Ala 245 250 255
- Val Thr Gln Leu Asn Leu Pro Asn Val Ala Met Tyr Leu Asp Ala Gly 260 265 270
- His Ala Gly Trp Leu Gly Trp Pro Ala Asn Gln Asp Pro Ala Ala Gln 275 280 285

Leu Phe Ala Asn Val Tyr Lys Asn Ala Ser Ser Pro Arg Ala Leu Arg

Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Gly Trp Asn Ile Thr Ser

Pro Pro Ser Tyr Thr Gln Gly Asn Ala Val Tyr Asn Glu Lys Leu Tyr

Ile His Ala Ile Gly Pro Leu Leu Ala Asn His Gly Trp Ser Asn Ala

Phe Phe Ile Thr Asp Gln Gly Arg Ser Gly Lys Gln Pro Thr Gly Gln

Gin Gin Trp Gly Asp Trp Cys Asn Val Ile Gly Thr Gly Phe Gly Ile

Arg Pro Ser Ala Asn Thr Gly Asp Ser Leu Leu Asp Ser Phe Val Trp 385 390

Val Lys Pro Gly Gly Glu Cys Asp Gly Thr Ser Asp Ser Ser Ala Pro 405

Arg Phe Asp Ser His Cys Ala Leu Pro Asp Ala Leu Gln Pro Ala Pro 420 425

Gln Ala Gly Ala Trp Phe Gln Ala Tyr Phe Val Gln Leu Leu Thr Asn 440

Ala Asn Pro Ser Phe Leu 450

<210> 63 <211> 1317 <212> DNA <213> Trichoderma reesei

<220>

<221> CDS

<222> (1)..(1317)

<223> Trichoderma reesei endoglucanase I

atg cag caa ccg gga acc agc acc ccc gag gtc cat ccc aag ttg aca Met Gln Gln Pro Gly Thr Ser Thr Pro Glu Val His Pro Lys Leu Thr

Page 86

1			5					10					15		
acc ta Thr Ty	ic aag /r Lys	tgc Cys 20	aca Thr	aag Lys	tcc Ser	ggg Gly	999 Gly 25	t gc Cys	gtg Val	gcc Ala	cag Gln	gac Asp 30	acc Thr	tcg Ser	96
gto gt Val Va	c ctt al Leu 35	gac Asp	tgg Trp	aac Asn	tac Tyr	cgc Arg 40	tgg Trp	atg Met	cac His	gac Asp	gca Ala 45	aac Asn	tac Tyr	aac Asn	144
tog tog Ser Cy 50	ys Thr	gtc Val	aac Asn	ggc Gly	ggc Gly 55	gtc Val	aac Asn	acc Thr	acg Thr	ctc Leu 60	t gc Cys	cct Pro	gac Asp	gag Glu	192
gcg ac Ala Th	ed tgt hr Cys	ggc Gly	aag Lys	aac Asn 70	tgc Cys	t t c Phe	atc Ile	gag Glu	ggc Gly 75	gtc Val	gac Asp	tac Tyr	gcc Ala	gcc Ala 80	240
tog go Ser Gl	gc gtc ly Val	acg Thr	acc Thr 85	tcg Ser	ggc Gly	agc Ser	agc Ser	ctc Leu 90	acc Thr	atg Met	aac Asn	cag Gln	tac Tyr 95	atg Met	288
ccc aq Pro Se	gc agc er Ser	tct Ser 100	ggc Gly	ggc Gly	tac Tyr	agc Ser	agc Ser 105	gtc Val	tct Ser	cct Pro	cgg Arg	ctg Leu 110	tat Tyr	ctc Leu	336
ctg ga Leu As	ac tct sp Ser 115	Asp	ggt Gly	gag Glu	tac Tyr	gtg Val 120	atg Met	ctg Leu	aag Lys	ct¢ Leu	aac Asn 125	ggc Gly	cag Gln	gag Glu	384
Leu Se	go tto er Phe 30	gac Asp	gtc Val	gac Asp	ctc Leu 135	tct Ser	gct Ala	ct g Leu	ccg Pro	tgt Cys 140	GIA	gag Glu	aac Asn	ggc Gly	432
tog of Ser Lo	to tac eu Tyr	ctg Leu	tct Ser	cag Gln 150	atg Met	gac Asp	gag Glu	aac Asn	ggg Gly 155	GIÀ	gcc Ala	aac Asn	cag Gln	tat Tyr 160	480
aac a Asn Ti	eg ged hr Ala	: ggt i Gly	gcc Ala 165	Asn	tac Tyr	999 Gly	agc Ser	ggc Gly 170	IAT	tgc Cys	gat Asp	gct Ala	cag Gln 175	0,0	528
ccc g Pro V	to caq al Glr	g aca n Thr 180	Trp	agg Arg	aac Asn	ggc Gly	acc Thr 185	ren	aac Asn	act Thr	agc Ser	cac His 190	0111	ggc Gly	576
tic to Phe C	ge tge ys Cys 195	s Asn	gag Glu	atg Met	gat Asp	atc Ile 200	Leu	gag Glu	ggc	aac Asn	tcg Ser 205	9	gcg Ala	aat Asn	624
Ala L	tg acc eu Thi	cct Pro	cac His	tct Ser	tgc Cys 215	Thr	gcc	acg Thr	gcc	Cys 220	, Asp	tct Ser	gcc Ala	ggt Gly	672
tạc g Cys G 225	gc tte Sly Phe	c aac ∋ Asr	ccc Pro	tat Tyr 230	Gly	agc Ser	ggc Gly	tac Tyr	Lys 235	. ser	tac Tyr	tac Tyr	ggc	Pro 240	720

gga Gly	gat Asp	acc Thr	gtt Val	gac Asp 245	acc Thr	tcc Ser	aag Lys	acc Thr	ttc Phe 250	acc Thr	atc Ile	atc Ile	acc Thr	cag Gln 255	ttc Phe	768
aac Asn	acg Thr	gac Asp	aac Asn 260	ggc Gly	t cg Ser	ccc Pro	tog Ser	ggc Gly 265	aac Asn	ctt Leu	gtg Val	agc Ser	atc Ile 270	acc Thr	cgc Arg	816
aag Lys	tac Tyr	cag Gln 275	caa Gln	aac Asn	ggc Gly	gtc Val	gac Asp 280	atc Ile	ccc Pro	agc Ser	gcc Ala	cag Gln 285	ccc Pro	ggc Gly	ggc Gly	864
gac Asp	acc Thr 290	atc lle	tog Ser	tcc Ser	t gc Cys	ccg Pro 295	tcc Ser	gcc Ala	tca Ser	gcc Ala	tac Tyr 300	ggc Gly	ggc Gly	ctc Leu	gcc Ala	912
acc Thr 305	atg Met	ggc Gly	aag Lys	gcc Ala	ctg Leu 310	agc Ser	agc Ser	ggc Gly	atg Met	gtg Val 315	ctc Leu	gtg Val	ttc Phe	agc Ser	att Ile 320	960
tgg Trp	aac Asn	gac Asp	aac Asn	agc Ser 325	cag Gln	tac Tyr	atg Met	aac Asn	tgg Trp 330	ctc Leu	gac Asp	agc Ser	ggc Gly	aac Asn 335	gcc Ala	1008
ggc Gly	ccc Pro	tgc Cys	agc Ser 340	agc Ser	acc Thr	gag Glu	ggc Gly	aac Asn 345	cca Pro	tcc Ser	aac Asn	acc Thr	ctg Leu 350	gcc Ala	aac Asn	1056
aac Asn	ccc Pro	aac Asn 355	acg Thr	cac His	gtc Val	gtc Val	ttc Phe 360	tcc Ser	aac Asn	atc Ile	cgc Arg	tgg Trp 365	gga Gly	gac Asp	att Ile	1104
999 Gly	tct Ser 370	act Thr	acg Thr	aac Asn	tcg Ser	act Thr 375	gcg Ala	ccc Pro	ccg Pro	ccc Pro	ccg Pro 380	cct Pro	gcg Ala	tcc Ser	agc Ser	1152
acg Thr 385	acg Thr	ttt Phe	t cg Se r	act Thr	aca Thr 390	cgg Arg	agg Arg	agc Ser	t c g Se r	acg Thr 395	act Thr	t cg Ser	agc Ser	agc Ser	ccg Pro 400	1200
agc Ser	t gc Cys	acg Thr	cag Gln	act Thr 405	His	tgg Trp	ggg Gly	cag Gln	tgc Cys 410	ggt Gly	ggc Gly	att Ile	ggg Gly	tac Tyr 415	agc Ser	1248
ggg Gly	t gc Cys	aag Lys	acg Thr 420	Cys	acg Thr	tcg Ser	ggc	act Thr 425	acg Thr	tgc Cys	cag Gln	tat Tyr	agc Ser 430	MSII	gac Asp	1296
tac Tyr	tac Tyr	tog Ser 435	Gln	t gc Cys	ctt Leu	tag										1317
<21 <21 <21 <21	1 > 2 >	84 438 PRT Tric	hode	rma	rees	ei										

<400> 84

Met Gln Gln Pro Gly Thr Ser Thr Pro Glu Val His Pro Lys Leu Thr l 5 10

Thr Tyr Lys Cys Thr Lys Ser Gly Gly Cys Val Ala Gln Asp Thr Ser 25 30

Val Val Leu Asp Trp Asn Tyr Arg Trp Met His Asp Ala Asn Tyr Asn 35 40 45

Ser Cys Thr Val Asn Gly Gly Val Asn Thr Thr Leu Cys Pro Asp Glu 50 55 60

Ala Thr Cys Gly Lys Asn Cys Phe Ile Glu Gly Val Asp Tyr Ala Ala 65 70 75 80

Ser Gly Val Thr Thr Ser Gly Ser Ser Leu Thr Met Asn Gln Tyr Met 85 90 95

Fro Ser Ser Ser Gly Gly Tyr Ser Ser Val Ser Pro Arg Leu Tyr Leu 100 105 110

Leu Asp Ser Asp Gly Glu Tyr Val Met Leu Lys Leu Asn Gly Gln Glu 115 120 125

Leu Ser Phe Asp Val Asp Leu Ser Ala Leu Pro Cys Gly Glu Asn Gly 130 135

Ser Leu Tyr Leu Ser Gln Met Asp Glu Asn Gly Gly Ala Asn Gln Tyr 145 155 160

Asn Thr Ala Gly Ala Asn Tyr Gly Ser Gly Tyr Cys Asp Ala Gln Cys 165 170 175

Pro Val Gln Thr Trp Arg Asn Gly Thr Leu Asn Thr Ser His Gln Gly 180 185 190

Phe Cys Cys Asn Glu Met Asp Ile Leu Glu Gly Asn Ser Arg Ala Asn 195 200 205

Ala Leu Thr Pro His Ser Cys Thr Ala Thr Ala Cys Asp Ser Ala Gly 210 215 220

Cys Gly Phe Asn Pro Tyr Gly Ser Gly Tyr Lys Ser Tyr Tyr Gly Pro 225 230 235 240

Gly Asp Thr Val Asp Thr Ser Lys Thr Phe Thr Ile Ile Thr Gln Phe 245 250 255

Asn Thr Asp Asn Gly Ser Pro Ser Gly Asn Leu Val Ser Ile Thr Arg 260 265 270

Lys Tyr Gln Gln Asn Gly Val Asp Ile Pro Ser Ala Gln Pro Gly Gly 275 280 285

Asp Thr Ile Ser Ser Cys Pro Ser Ala Ser Ala Tyr Gly Gly Leu Ala 290 300

Thr Met Gly Lys Ala Leu Ser Ser Gly Met Val Leu Val Phe Ser Ile 305 310 315

Trp Asn Asp Asn Ser Gln Tyr Met Asn Trp Leu Asp Ser Gly Asn Ala 325 330 335

Gly Pro Cys Ser Ser Thr Glu Gly Asn Pro Ser Asn Thr Leu Ala Asn 340 345 350

Asn Pro Asn Thr His Val Val Phe Ser Asn Ile Arg Trp Gly Asp Ile 355 360 365

Gly Ser Thr Thr Asn Ser Thr Ala Pro Pro Pro Pro Ala Ser Ser 370 380

Thr Thr Phe Ser Thr Thr Arg Arg Ser Ser Thr Thr Ser Ser Ser Pro 385 390 395 400

Ser Cys Thr Gln Thr His Trp Gly Gln Cys Gly Gly Ile Gly Tyr Ser 405 410 415

Gly Cys Lys Thr Cys Thr Ser Gly Thr Thr Cys Gln Tyr Ser Asn Asp 420 430

Tyr Tyr Ser Gln Cys Leu

<210> 85 <211> 954

Page 90

```
<212> DNA
<2:3> Artificial Sequence
<220>
<223> 6GP1
<220>
<221> CDS
<222> (1)..(954)
<223> 6GP1
<400> 85
atg ggc gtg gac ccg ttc gag cgc aac aag atc ctc ggc cgc ggc atc
                                                                       48
Met Gly Val Asp Pro Phe Glu Arg Asn Lys Ile Leu Gly Arg Gly Ile
                                                                       96
aac atc ggc aac gcc ctg gag gcc ccg aac gag ggc gac tgg ggc gtg
Asn Ile Gly Asn Ala Leu Glu Ala Pro Asn Glu Gly Asp Trp Gly Val
                                                                      144
gtg atc aag gac gag ttc ttc gac atc atc aag gag gcc ggc ttc tcc
Val Ile Lys Asp Glu Phe Phe Asp Ile Ile Lys Glu Ala Gly Phe Ser
                            40
                                                                      192
cac gtg cgc atc ccg atc cgc tgg tcc acc cac gcc tac gcc ttc ccg
His Val Arg Ile Pro Ile Arg Trp Ser Thr His Ala Tyr Ala Phe Pro
                        55
                                                                      240
ccg tac aag atc atg gac cgc ttc ttc aag cgc gtg gac gag gtg atc
Pro Tyr Lys Ile Met Asp Arg Phe Phe Lys Arg Val Asp Glu Val Ile
aac ggc gcc ctc aag cgc ggc ctc gcc gtg gcc atc aac atc cac cac
                                                                      288
Asn Gly Ala Leu Lys Arg Gly Leu Ala Val Ala Ile Asn Ile His His
tac gag gag ctc atg aac gac ccg gag gag cac aag gag cgc ttc ctc
                                                                       336
Tyr Glu Glu Leu Met Asn Asp Pro Glu Glu His Lys Glu Arg Phe Leu
                                105
                                                                       384
qcc ctc tqq aag cag atc gcc gac cgc tac aag gac tac ccg gag acc
Ala Leu Trp Lys Gln Ile Ala Asp Arg Tyr Lys Asp Tyr Pro Glu Thr
                                                                       432
ctc ttc ttc gag atc ctc aac gag ccg cac ggc aac ctc acc ccg gag
Leu Phe Phe Glu Ile Leu Asn Glu Pro His Gly Asn Leu Thr Pro Glu
   130
aag tgg aac gag etg etc gag gag gee etc aag gtg ate ege tee ate
                                                                      480
Lys Trp Asn Glu Leu Leu Glu Glu Ala Leu Lys Val Ile Arg Ser Ile
gac aag aag cac acc atc atc att ggc acc gca gag tgg gga ggc atc
                                                                      528
Asp Lys Lys His Thr Ile Ile Ile Gly Thr Ala Glu Trp Gly Gly Ile
                165
tcc gcc ctc gag aag ctc tcc gtg ccg aag tgg gag aag aat tcc atc
                                                                      576
```

Page 91

Ser Ala Leu Glu Lys Leu Ser Val Pro Lys Trp Glu Lys Asn-Ser Ile 180 185 190	
gtg acc atc cac tac tac aac ccg ttc gag ttc acg cac cag ggc gcc Val Thr !le His Tyr Tyr Asn Pro Phe Glu Phe Thr His Gln Gly Ala 195 200 205	624
gag tgg gtg gag ggc tcc gag aag tgg ctt ggc cgc aag tgg ggc tcc Glu Trp Val Glu Gly Ser Glu Lys Trp Leu Gly Arg Lys Trp Gly Ser 210 215 220	672
ccg çac çac cag aag cac ctc atc gag gag ttc aac ttc atc gag gag Pro Asp Asp Gln Lys His Leu Ile Glu Glu Phe Asn Phe Ile Glu Glu 235 230 235 240	720
tgg too aag aag aac aag ogo oog ato tac ato ggo gag ttt ggo goo Trp Ser Lys Lys Asn Lys Arg Pro Ile Tyr Ile Gly Glu Phe Gly Ala 245 250 255	768
tac cgc aag gcc gac ctc gag tcc cgc atc aag tgg acc tcc ttc gtg Tyr Arg Lys Ala Asp Leu Glu Ser Arg Ile Lys Trp Thr Ser Phe Val 260 265 270	816
gtg cgt gag atg gag aag cgc cgc tgg tcc tgg gcc tac tgg gag ttc Val Arg Glu Met Glu Lys Arg Arg Trp Ser Trp Ala Tyr Trp Glu Phe 275 280 285	864
tgc tcc ggc ttc ggc gtg tac gac acc ctc cgc aag acc tgg aac aag Cys Ser Gly Phe Gly Val Tyr Asp Thr Leu Arg Lys Thr Trp Asn Lys 290 295 300	912
gac etc etc gag gec etc atc gge gge gac tec atc gag tag Asp Leu Glu Ala Leu Ile Gly Gly Asp Ser Ile Glu 305 310	954
<210> 86 <211> 317 <212> PRT <213> Artificial Sequence	
<220> <223> Synthetic Construct	
<400> 86	
Net Gly Val Asp Pro Phe Glu Arg Asn Lys Ile Leu Gly Arg Gly Ile 1 10 15	
Asn Ile Gly Asn Ala Leu Glu Ala Pro Asn Glu Gly Asp Trp Gly Val 20 25 30	

Page 92

Val Ile Lys Asp Glu Phe Phe Asp Ile Ile Lys Glu Ala Gly Phe Ser 35 40 45

His	Vā l	Arg	lle	Pro	Ile	Arg	Trp	Ser	Thr	His	Ala	Tyr	Ala	Phe	Pro
	50	•				55					60				

- Pro Tyr Lys Ile Met Asp Arg Phe Phe Lys Arg Val Asp Glu Val Ile 65 70 75 80
- Asn Gly Ala Leu Lys Arg Gly Leu Ala Val Ala Ile Asn Ile His His 85 90 95
- Tyr Glu Glu Leu Met Asn Asp Pro Glu Glu His Lys Glu Arg Phe Leu 100 105 110
- Ala Leu Trp Lys Gln Ile Ala Asp Arg Tyr Lys Asp Tyr Pro Glu Thr
- Leu Phe Phe Glu Ile Leu Asn Glu Pro His Gly Asn Leu Thr Pro Glu 130 140
- Lys Trp Asn Glu Leu Leu Glu Glu Ala Leu Lys Val Ile Arg Ser Ile 145 150 155 160
- Asp Lys Lys His Thr Ile Ile Gly Thr Ala Glu Trp Gly Gly Ile 165 170 175
- Ser Ala Leu Glu Lys Leu Ser Val Pro Lys Trp Glu Lys Asn Ser Ile 180 185 190
- Val Thr Ile His Tyr Tyr Asn Pro Phe Glu Phe Thr His Gln Gly Ala
- Glu Trp Val Glu Gly Ser Glu Lys Trp Leu Gly Arg Lys Trp Gly Ser 210 220
- Pro Asp Asp Gln Lys His Leu Ile Glu Glu Phe Asn Phe Ile Glu Glu 225 230 240
- Trp Ser Lys Lys Asn Lys Arg Fro Ile Tyr Ile Gly Glu Phe Gly Ala 245 250 255
- Tyr Arg Lys Ala Asp Leu Glu Ser Arg Ile Lys Trp Thr Ser Phe Val 260 265 270
- Val Arg Glu Met Glu Lys Arg Arg Trp Ser Trp Ala Tyr Trp Glu Phe 275 280 285

Page 93

Cys Ser Gly 290	Phe Gly Val	Tyr Asp Thr 295	Leu Arg	Lys Thr Tr	o Asn Lys	
Asp Leu Leu 305	Glu Ala Leu 310	ile Gly Gly	Asp Ser 315	Ile Glu		
<210> 87 <211> 1248 <212> DNA <213> Horde	eum vulgare					
	.(1248) ey AmyI amyl	ase				
<400> 87 atg gca cac Met Ala His 1	caa gtc ctc Gln Val Leu 5	ttt cag ggg Phe Gln Gly	ttc aac Phe Asn 10	tgg gag tce Trp Glu Se	g tgg aag Trp Lys 15	48
cag age gge Gln Ser Gly	ggg tgg tac Gly Trp Tyr 20	aac atg atg Asn Met Met 25	atg ggc Met Gly	aag gtc gad Lys Val Asp 30	gac atc Asp Ile	96
gcc gct gcc Ala Ala Ala 35	gga gtc acc Gly Val Thr	cac gtc tgg His Val Trp 40	ctg cca Leu Pro	ccg ccg tcc Pro Pro Se 45	g cac tcc r His Ser	144
gtc tcc aac Val Ser Asn 50	gaa ggt tac Glu Gly Tyr	atg cct ggt Met Pro Gly 55	cgg ctg Arg Leu	tac gac ate Tyr Asp Ile 60	gac gcg Asp Ala	192
tcc aag tac Ser Lys Tyr 65	ggc aac gcg Gly Asn Ala 70	gcg gag ctc Ala Glu Leu	aag tcg Lys Ser 75	ctc atc gge Leu Ile Gl	gcg ctc Ala Leu 80	240
cac ggc aag His Gly Lys	ggc gtg cag Gly Val Gln 85	gcc atc gcc Ala Ile Ala	gac atc Asp Ile 90	gtc atc aac Val Ile Asi	cac cgc His Arg 95	288
tgc gcc gac Cys Ala Asp	tac aag gat Tyr Lys Asp 100	agc cgc ggc Ser Arg Gly 105	Ile Tyr	tgc atc tto Cys Ile Phe 110	Glu Gly	336
ggc acc tcc Gly Thr Ser 115	gac ggc cgc Asp Gly Arg	ctc gac tgg Leu Asp Trp 120	ggc ccc Gly Pro	cac atg ato His Met Ile 125	tgt cgc Cys Arg	384
gac gac acc Asp Asp Thr 130	aaa tac tcc Lys Tyr Ser	gat ggc acc Asp Gly Thr 135	gca aac Ala Asn	ctc gac acc Leu Asp Thi 140	gga gcc Gly Ala	432

Page 94

	Phe					Asp						gac Asp				480
					Trp							gac Asp				528
									Gly			ccg Pro				5,76
			lle									gtg Val 205				624
		Asn										aac Asn				672
												aag Lys				720
												aaa Lys				768
												gac Asp				816
•	•				_				-	-	•	gcc Ala 285	-			864
												tgg Trp				912
												acc Thr				960
		•				•						ggg Gly		•	-	1008
												ggc Gly				1056
acg Th <i>r</i>	agc Ser	gct Ala 355	ctg Leu	aag Lys	atc Ile	ctc Leu	atg Met 360	cac His	gaa Glu	gga Gly	gat Asp	gcc Ala 365	tac Tyr	gtc Val	gcc Ala	1104
												agg Arg				1152

370	375		380	
ggg gcg gtg atc Gi _/ Ala Val Ile 385	ccg gcc ggg Pro Ala Gly 390	ttc gtg acc Phe Val Thr	tcg gca cac ggc Ser Ala His Gly 395	aac gac 1200 Asn Asp 400
tac gcc gtc tgg Tyr Ala Val Trp	gag aag aac Glu Lys Asn 405	ggt gcc gcg Gly Ala Ala 410	gca aca cta caa Ala Thr Leu Gln	cgg agc 1248 Arg Ser 415
<210> 88 <211> 416 <212> PRT <213> Hordeum	vulgare			
<400> 88				
Met Ala His Gln l	Val Leu Phe 5	Gln Gly Phe 10	Asn Trp Glu Ser	Trp Lys 15
Gln Ser Gly Gly 20	Trp Tyr Asn	Met Met Met 25	Gly Lys Val Asp 30	Asp Ile
Ala Ala Ala Gly 35	Val Thr His	Val Trp Leu 40	Pro Pro Pro Ser 45	His Ser
Val Ser Asn Glu 50	Gly Tyr Met 55	Pro Gly Arg	Leu Tyr Asp Ile 60	Asp Ala
Ser Lys Tyr Gly 65	Asn Ala Ala 70	Glu Leu Lys	Ser Leu Ile Gly 75	Ala Leu 80
His Gly Lys Gly	Val Gln Ala 85	Ile Ala Asp 90	Ile Val Ile Asn	His Arg 95
Cys Ala Asp Tyr 100	Lys Asp Ser	Arg Gly Ile 105	Tyr Cys Ile Phe	
Gly Thr Ser Asp 115	Gly Arg Leu	Asp Trp Gly 120	Pro His Met Ile 125	Cys Arg
Asp Asp Thr Lys 130	Tyr Ser Asp 135	Gly Thr Ala	Asn Leu Asp Thr 140	Gly Ala
Aso Phe Ala Ala 145	Ala Pro Asp 150	Ile Asp His	Leu Asn Asp Arg 155	Val Gln 160

Arg Glu Leu Lys Glu Trp Leu Leu Trp Leu Lys Ser Asp Leu Gly Phe 165 170 175

Asp Ala Trp Arg Leu Asp Phe Ala Arg Gly Tyr Ser Pro Glu Met Ala 180 185 190

Lys Val Tyr Ile Asp Gly Thr Ser Pro Ser Leu Ala Val Ala Glu Val 195 200 205

Trp Asp Asn Met Ala Thr Gly Gly Asp Gly Lys Pro Asn Tyr Asp Gln 210 215 220

Asp Ala His Arg Gln Asn Leu Val Asn Trp Val Asp Lys Val Gly Gly 225 230 235 240

Ala Ala Ser Ala Gly Met Val Phe Asp Phe Thr Thr Lys Gly Ile Leu 245 250 255

Asn Ala Ala Val Glu Gly Glu Leu Trp Arg Leu Ile Asp Pro Gln Gly 260 265 270

Lys Ala Pro Gly Val Met Gly Trp Trp Pro Ala Lys Ala Val Thr Phe 275 280 285

Val Asp Asn His Asp Thr Gly Ser Thr Gln Ala Met Trp Pro Phe Pro 290 295 300

Ser Asp Lys Val Met Gln Gly Tyr Ala Tyr Ile Leu Thr His Pro Gly 305 310 315

Ile Pro Cys Ile Phe Tyr Asp His Phe Phe Asn Trp Gly Phe Lys Asp 325 330 335

Gln Ile Ala Ala Leu Val Ala Ile Arg Lys Arg Asn Gly Ile Thr Ala 340 345 350

Thr Ser Ala Leu Lys Ile Leu Met His Glu Gly Asp Ala Tyr Val Ala 355 360 365

Glu Ile Asp Gly Lys Val Val Lys Ile Gly Ser Arg Tyr Asp Val 370 380

Gly Ala Val Ile Pro Ala Gly Phe Val Thr Ser Ala His Gly Asn Asp 385 390 395 400

Page 97

Tyr Ala Val Trp Glu Lys Asn Gly Ala Ala Ala Thr Leu Gln Arg Ser 405 410 415	
<210> 89 <211> 1401 <212> DNA <213> Artificial Sequence	
<220> <223> Trichoderma reesei ß-Glucosidase 2	
<220> <221> CDS <222> (1)(1401) <223> Trichoderma reesei ß-Glucosidase 2	
<pre><400> 89 atg ttg ccc aag gac ttt cag tgg ggg ttc gcc acg gct gcc tac cag Met Leu Pro Lys Asp Phe Gln Trp Gly Phe Ala Thr Ala Ala Tyr Gln 1 5 10 15</pre>	48
atc gag ggc gcc gtc gac cag gac ggc cgc ggc ccc agc atc tgg gac Ile Glu Gly Ala Val Asp Gln Asp Gly Arg Gly Pro Ser Ile Trp Asp 20 25 30	96
acg ttc tgc gcg cag ccc ggc aag atc gcc gac ggc tcg tcg ggc gtg Thr Phe Cys Ala Gln Pro Gly Lys Ile Ala Asp Gly Ser Ser Gly Val 35 40 45	144
acg gcg tgc gac tcg tac aac cgc acg gcc gag gac att gcg ctg ctg Thr Ala Cys Asp Ser Tyr Asn Arg Thr Ala Glu Asp Ile Ala Leu Leu 50 55 60	192
aag tog oto ggg goo aag ago tao ogo tto too ato tog tgg tog ogo Lys Ser Leu Gly Ala Lys Ser Tyr Arg Phe Ser Ile Ser Trp Ser Arg 65 70 75 80	240
atc atc ccc gag ggc ggc cgc ggc gat gcc gtc aac cag gcg ggc atc Ile Ile Pro Glu Gly Gly Arg Gly Asp Ala Val Asn Gln Ala Gly Ile 85 90 95	288
gac cac tac gtc aag ttc gtc gac gac ctg ctc gac gcc ggc atc acg Asp His Tyr Val Lys Phe Val Asp Asp Leu Leu Asp Ala Gly Ile Thr 100 105 110	336
Pro Phe Ile Thr Leu Phe His Trp Asp Leu Pro Glu Gly Leu His Gln 115	384
cgg tac ggg ggg ctg ctg aac cgc acc gag ttc ccg ctc gac ttt gaa Arg Tyr Gly Gly Leu Leu Asn Arg Thr Glu Phe Pro Leu Asp Phe Glu 130 135 . 140	432
aac tac gcc cgc gtc atg ttc agg gcg ctg ccc aag gtg cgc aac tgg	480

Page 98

Asn 145		Ala	Arg	Val	Met 150	Phe	Arg	Ala	Leu	Pro 155	Lys	Val	Arg	Asn	Trp 160	
												ggc Gly				528
												ccg Pro				576
												aag Lys 205				624
gac Asp	gac Asp 210	ttc Phe	aag Lys	ccc Pro	gcc Ala	agc Ser 215	ggc Gly	gac Asp	ggc Gly	cag Gln	atc Ile 220	ggc Gly	atc Ile	gtc .Val	ctc Leu	672
												ccg Pro				720
gag Glu	gcg Ala	gcc Ala	gag Glu	cgg Arg 245	cgc Arg	ctc Leu	gag Glu	ttc Phe	ttc Phe 250	acg Thr	gcc Ala	tgg Trp	ttc Phe	gcg Ala 255	gac Asp	768
ccc Pro	ātc Ile	tac Tyr	ttg Leu 260	ggc Gly	gac Asp	tac Tyr	ccg Pro	gcg Ala 265	t cg Ser	atg Met	cgc Arg	aag Lys	cag Gln 270	ctg Leu	ggc Gly	816
gac Asp	cgg Arg	ctg Leu 275	ccg Pro	acc Thr	ttt Phe	acg Thr	ccc Pro 280	gag Glu	gag Glu	cgc Arg	gcc Ala	ctc Leu 285	gtc Val	cac His	ggc Gly	864
tcc Ser	aac Asn 290	gac Asp	ttt Phe	tac Tyr	ggc Gly	atg Met 295	aac Asn	cac His	tac Tyr	acg Thr	tcc Ser 300	aac Asn	tac Tyr	atc Ile	cgc Arg	912
cac His 305	cgc Arg	agc Ser	t cg Ser	ccc Pro	gcc Ala 310	tcc Ser	gcc Ala	gac Asp	gac Asp	acc Thr 315	gtc Val	ggc Gly	aac Asn	gtc Val	gac Asp 320	960
gtg Val	ctc Leu	ttc Phe	acc Thr	aac Asn 325	aag Lys	cag Gln	ggc Gly	aac Asn	tgc Cys 330	atc Ile	ggc Gly	ccc Pro	gag Glu	acg Thr 335	cag Gln	1008
tcc Ser	ccc Pro	tgg Trp	ctg Leu 340	cgc Arg	ccc Pro	tgt Cys	gcc Ala	gcc Ala 345	ggc Gly	ttc Phe	cgc Arg	gac Asp	ttc Phe 350	ctg Leu	gtg Val	1056
tgg Trp	atc Ile	agc Ser 355	aag Lys	agg Arg	tac Tyr	ggc Gly	tac Tyr 360	ccg Pro	ccc Pro	atc Ile	tac Tyr	gtg Val 365	acg Thr	gag Glu	aac Asn	1104
ggc Gly	acg Thr 370	agc Ser	atc Ile	aag Lys	ggc Gly	gag Glu 375	agc Ser	gac Asp	t t g Leu	ccc Pro	aag Lys 380	gag Glu	aag Lys	att Ile	ctc Leu	1152

Page 99

G1u 385	gat Asp	gac Asp	ttc Phe	agg Arg	gtc Val 390	aag Lys	tac Tyr	tat Tyr	aac Asn	gag Glu 395	tac Tyr	atc Ile	cgt Arg	gcc Ala	atg Met 400	1200
gtt Val	acc Thr	gcc Ala	gtg Val	gag Glu 405	ctg Leu	gac Asp	ggg Gly	gtc Val	aac Asn 410	gtc Val	aag Lys	999 Gly	tac Tyr	ttt Phe 415	gcc Ala	1248
tgg Trp	tcg Ser	ctc Leu	atg Met 420	gac Asp	aac Asn	ttt Phe	gag Glu	tgg Trp 425	gcg Ala	gac Asp	ggc Gly	tac Tyr	gtg Val 430	acg Thr	agg Arg	1296
ttt Pne	G] À 333	gtt Val 435	acg Thr	tat Tyr	gtg Val	gat Asp	tat Tyr 440	gag Glu	aat Asn	999 Gly	cag Gln	aag Lys 445	cgg Arg	ttc Phe	ccc Pro	1344
aag Lys	aag Lys 450	agc Ser	gca Ala	aag Lys	agc Ser	ttg Leu 455	aag Lys	ccg Pro	ctg Leu	ttt Phe	gac Asp 460	gag Glu	ctg Leu	att Ile	gcg Ala	1392
	gcg Ala	tga														1401
<219 <211 <211 <211	1 > 4 2 > 1		ficia	al Se	equer	nce										
					-											
<220 <22		Synth	netio	: Cor	istri	ıct										
	3> \$	Synth 30	netio	: Cor	nstri	ıct										
< 400	3> \$ 0> \$	90	netic Lys				Trp	Gly	Phe 10	Ala	Thr	Ala	Ala	Tyr 15	Gln	
<223 <400 Met 1	3>	90 Pro		Asp 5	Phe	Gln			10					15		
<22: <400 Met 1	3> 5 Leu Glu	Pro Gly	Lys	Asp 5 Val	Phe Asp	Gln Gln	Asp	Gly 25	10 Arg	Gly	Pro	Ser	Ile 30	15 Trp	Asp	
<22: <400 Met 1 Ile	3> S Leu Glu Phe	Pro Gly Cys 35	Lys Ala 20	Asp 5 Val Gln	Phe Asp Pro	Gln Gln Gly	Asp Lys 40	Gly 25	Arg Ala	Gly Asp	Pro Gly	Ser Ser 45	Ile 30 Ser	Trp	Asp Val	
<22: <400 Met 1 !le Thr	Glu Phe Ala	Pro Gly Cys 35	Lys Ala 20	Asp 5 Val Gln Ser	Phe Asp Pro	Gln Gln Gly Asn 55	Asp Lys 40 Arg	Gly 25 Ile Thr	Arg Ala	Gly Asp Glu	Pro Gly Asp 60	Ser Ser 45	Ile 30 Ser	Trp Gly Leu	Asp Val Leu	

Page 100

Asp	His	Туг	Val 100	Lys	Phe	Val	Asp	Asp 105	Leu	Leu	Asp	Ala	Gly 110	Ile	Thr
Pro	Phe	Ile 115	Thr	Leu	Phe	His	Trp 120	Asp	Leu	Pro	Glu	Gly 125	Leu	His	Gln
Arg	Tyr 130	Gly	Gly	Leu	Leu	Asn 135	Arg	Thr	Glu	Phe	Pro 140	Leu	Asp	Phe	Glu
Asn 145	Tyr	Ala	Arg	Val	Met 150	Phe	Arg	Ala	Leu	Pro 155	Lys	Val	Arg	Asn	Trp 160
1 1 e	Thr	Phe	A.s n	Glu 165	Pro	Leu	Cys	Ser	Ala 170	Ile	Pro	Gly	Tyr	Gly 175	Ser
Gly	Thr	Phe	Ala 180	Pro	Gly	Arg	Gln	Ser 185	Thr	Ser	Glu	Pro	Trp 190	Thr	Val
Gly	His	Asn 195	Ile	Leu	Val	Ala	His 200	Gly	Arg	Ala	Val	Lys 205	Ala	Tyr	Arg
Asp	Asp 210	Phe	Lys	Pro	Ala	Ser 215	Gly	Asp	Gly	Gln	11e 220	Gly	Ile	Val	Leu
Asn 225	Gly	Asp	Phe	Thr	Tyr 230	Pro	Trp	Asp	Ala	Ala 235	Asp	Pro	Ala	Asp	Lys 240
Glu	Ala	Ala	Glu	Arg 245	Arg	Leu	Glu	Phe	Phe 250	Thr	Ala	Trp	Phe	Ala 255	Asp
Pro	Ile	Tyr	Leu 260	Gly	Asp	Tyr	Pro	Ala 265	Ser	Met	Arg	Lys	Gln 270	Leu	Gly
Asp	Arg	Leu 275	Pro	Thr	Phe	Thr	Pro 280	Glu	Glu	Arg	Ala	Leu 285	Val	His	Gly
Ser	Asn 290	Asp	Phe	Tyr	Gly	Met 295	Asn	His	Tyr	Thr	Ser 300	Asn	Tyr	Ile	Arg
His 305	Arg	Ser	Ser	Pro	Ala 310	Ser	Ala	Asp	Asp	Thr 315	Val	Gly	Asn	Val	Asp 320

Page 101

Val Leu Phe Thr Asn Lys Gln Gly Asn Cys Ile Gly Pro Glu Thr Gln Ser Pro Trp Leu Arg Pro Cys Ala Ala Gly Phe Arg Asp Phe Leu Val Trp lle Ser Lys Arg Tyr Gly Tyr Pro Pro Ile Tyr Val Thr Glu Asn 360 365 355 Gly Thr Ser Ile Lys Gly Glu Ser Asp Leu Pro Lys Glu Lys Ile Leu 375 370 Glu Asp Asp Phe Arg Val Lys Tyr Tyr Asn Glu Tyr Ile Arg Ala Met 390 Val Thr Ala Val Glu Leu Asp Gly Val Asn Val Lys Gly Tyr Phe Ala 410 Trp Ser Leu Met Asp Asn Phe Glu Trp Ala Asp Gly Tyr Val Thr Arg 425 Phe Gly Val Thr Tyr Val Asp Tyr Glu Asn Gly Gln Lys Arg Phe Pro Lys Lys Ser Ala Lys Ser Leu Lys Pro Leu Phe Asp Glu Leu Ile Ala 455 Ala Ala 465 <210> 91 <211> 2103 <212> DNA <213> Artificial Sequence <223> Trichoderma reesei ß-Glucosidase D <220> <221> CDS <222> (1)..(2103) <223> Trichoderma reesei ß-Glucosidase D atg att ctc ggc tgt gaa agc aca ggt gtc atc tct gcc gtc aaa cac 48 Met Ile Leu Gly Cys Glu Ser Thr Gly Val Ile Ser Ala Val Lys His

ttt Phe	gtc Val	gcc Ala	aac Asn 20	gac Asp	cag Gln	gag Glu	cac His	gag Glu 25	cgg Arg	cga Arg	gcg Ala	gtc Val	gac Asp 30	tgt Cys	ctc Leu	96
atc Ile	acc Thr	cag Gln 35	cgg Arg	gct Ala	ctc Leu	cgg Arg	gag Glu 40	gtc Val	tat Tyr	ctg Leu	cga Arg	ccc Pro 45	ttc Phe	cag Gln	atc Ile	144
gta Val	gcc Ala 50	cga Arg	gat Asp	gca Ala	agg Arg	ccc Pro 55	ggc Gly	gca Ala	ttg Leu	atg Met	aca Thr 60	tcc Ser	tac Tyr	aac Asn	aag Lys	192
gtc Val 65	aat Asn	ggc Gly	aag Lys	cac His	gtc Val 70	gct Ala	gac Asp	agc Ser	gcc Ala	gag Glu 75	ttc Phe	ctt Leu	cag Gln	ggc Gly	att Ile 80	240
ctc Leu	cąg Arg	act Thr	gag Glu	tgg Trp 85	aat Asn	tgg Trp	gac Asp	cct Pro	ctc Leu 90	att Ile	gtc Val	agc Ser	gac Asp	tgg Trp 95	tac Tyr	288
ggc Gly	acc Thr	tac Tyr	acc Thr 100	act Thr	att Ile	gat Asp	gcc Ala	atc Ile 105	aaa Lys	gcc Ala	ggc Gly	ctt Leu	gat Asp 110	ctc Leu	gag Glu	336
atg Met	ccg Pro	ggc Gly 115	gtt Val	tca Ser	cga Arg	tat Tyr	cgc Arg 120	ggc Gly	aaa Lys	tac Tyr	atc Ile	gag Glu 125	tct Ser	gct Ala	ctg Leu	384
cag Gln	gcc Ala 130	cạt Arg	ttg Leu	ctg Leu	aag Lys	cag Gln 135	tcc Ser	act Thr	atc Ile	gat Asp	gag Glu 140	cgc Arg	gct Ala	cgc Arg	cgc Arg	432
gtg Val 145	ctc Leu	agg Arg	ttc Phe	gcc Ala	cag Gln 150	aag Lys	gcc Ala	agc Ser	cat His	ctc Leu 155	aag Lys	gtc Val	tcc Ser	gag Glu	gta Val 160	480
gag Glu	caa Gln	ggc Gly	cgt Arg	gac Asp 165	ttc Phe	cca Pro	gag Glu	gat Asp	cgc Arg 170	gtc Val	ctc Leu	aac Asn	cgt Arg	cag Gln 175	atc Ile	528
t gc Cys	ggc Gly	agc Ser	agc Ser 180	att Ile	gtc Val	cta Leu	ctg Leu	aag Lys 185	aat Asn	gag Glu	aac Asn	tcc Ser	atc Ile 190	tta Leu	cct Pro	576
ctc Leu	ccc Pro	aag Lys 195	tcc Ser	gtc Val	aag Lys	aag Lys	gtc Val 200	gcc Ala	ctt Leu	gtt Val	ggt Gly	tcc Ser 205	cac His	gtg Val	cgt Arg	624
cta Leu	ccg Pro 210	gct Ala	atc Ile	tcg Ser	gga Gly	gga Gly 215	ggc Gly	agc Ser	gcc Ala	tct Ser	ctt Leu 220	gtc Val	cct Pro	tac Tyr	tat Tyr	672
gcc Ala 225	ata Ile	tct Ser	cta Leu	tac Tyr	gat Asp 230	gcc Ala	gtc Val	tct Ser	gag Glu	gta Val 235	cta Leu	gcc Ala	ggt Gly	gcc Ala	acg Thr 240	720
ātc	acg	cac	gag	gtc	ggt	gcc	tat	gcc	cac	caa	atg	ctg	ccc	gtc	atc	768

Ile	Thr	His	Glu	Val 245	Gly	Ala	Tyr	Ala	His 250	Gln	Met	Leu	Pro	Val 255	Ile	
gac Asp	gca Ala	atg Met	atc Ile 260	agc Ser	aac Asn	gcc Ala	gta Val	atc Ile 265	cac His	ttc Phe	tac Tyr	aac Asn	gac Asp 270	ccc Pro	atc Ile	816
gat Asp	gtc Val	aaa Lys 275	gac Asp	aga Arg	aag Lys	ctc Leu	ctt Leu 280	ggc Gly	agt Ser	gag Glu	aac Asn	gta Val 285	tcg Ser	tog Ser	aca Thr	864
t og Ser	ttc Phe 290	cag Gln	ctc Leu	atg Net	gat Asp	tac Tyr 295	aac Asn	aac Asn	atc Ile	cca Pro	acg Thr 300	ctc Leu	aac Asn	aag Lys	gcc Ala	912
atq Met 305	ttc Phe	tgg Trp	ggt Gly	act Thr	ctc Leu 310	gtg Val	ggc Gly	gag Glu	ttt Phe	atc Ile 315	cct Pro	acc Thr	gcc Ala	acg Thr	gga Gly 320	960
att Ile	tgg Trp	gaa Glu	ttt Phe	ggc Gly 325	ctc Leu	agt Ser	gtc Val	ttt Phe	ggc Gly 330	act Thr	gcc Ala	gac Asp	ctt Leu	tat Tyr 335	att Ile	1008
gat Asp	aat Asn	gag Glu	ctc Leu 340	gtg Val	att Ilė	gaa Glu	aat Asn	aca Thr 345	aca Thr	cat His	cag Gln	acg Thr	cgt Arg 350	gga Gly	acc Thr	1056
gcc Ala	ttt Phe	ttc Phe 355	gga Gly	aag Lys	gga Gly	acg Thr	acg Thr 360	gaa Glu	aaa Lys	gtc Val	gct Ala	acc Thr 365	agg Arg	agg Arg	atg Met	1104
gtg Val	gcc Ala 370	ggc Gly	agc Ser	acc Thr	tac Tyr	aag Lys 375	ctg Leu	cgt Arg	ctc Leu	gag Glu	ttt Phe 380	Gly	tct Ser	gcc Ala	aac Asn	1152
acg Thr 385	acc Thr	aag Lys	atg Met	gag Glu	acg Thr 390	acc Thr	ggt Gly	gtt Val	gtc Val	aac Asn 395	ttt Phe	ggc Gly	ggc Gly	ggt Gly	gcc Ala 400	1200
gta Val	cac His	ctg Leu	ggt Gly	gcc Ala 405	tgt Cys	ctc Leu	aag Lys	gtc Val	gac Asp 410	cca Pro	cag Gln	gag Glu	atg Met	att Ile 415	gcg Ala	1248
cgg Arg	gcc Ala	gtc Val	aag Lys 420	gcc Ala	gca Ala	gcc Ala	gat Asp	gcc Ala 425	gac Asp	tac Tyr	acc Thr	atc Ile	atc Ile 430	tgc Cys	acg Thr	1296
gga Gly	ctc Leu	agc Ser 435	ggc Gly	gag Glu	tgg Trp	gag Glu	tct Ser 440	gag Glu	ggt Gly	ttt Phe	gac Asp	cgg Arg 445	cct Pro	cac His	atg Met	1344
gac Asp	ct g Leu 450	ccc Pro	cct Pro	ggt Gly	gtg Val	gac Asp 455	acc Thr	atg Met	atc Ile	t cg Ser	caa Gln 460	gtt Val	ctt Leu	gac Asp	gcc Ala	1392
gct Ala 465	ccc Pro	aat Asn	gct Ala	gta Val	gtc Val 470	gtc Val	aac Asn	cag Gln	tca Ser	ggc Gly 475	acc Thr	cca Pro	gtg Val	aca Thr	atg Met 480	1440

Page 104

agc Ser	tgg Trp	gct Ala	cat His	aaa Lys 485	gca Ala	aag Lys	gcc Ala	att Ile	gtg Val 490	cag Gln	gct Ala	tgg Trp	tat Tyr	ggt Gly 495	ggt Gly	1488
aac Asn	gag Glu	aca Thr	ggc Gly 500	cac His	gga Gly	atc Ile	tcc Ser	gat Asp 505	gtg Val	ctc Leu	ttt Phe	ggc Gly	aac Asn 510	gtc Val	aac Asn	1536
ccg Pro	tcg Ser	999 Gly 515	aaa Lys	ctc Leu	tcc Ser	cta Leu	tcg Ser 520	tgg Trp	cca Pro	gtc Val	gat Asp	gtg Val 525	aag Lys	cac His	aac Asn	1584
cca Pro	gca Ala 530	tat Tyr	ctc Leu	aac Asn	tac Tyr	gcc Ala 535	agc Ser	gtt Val	ggt Gly	gga Gly	cgg Arg 540	gtc Val	ttg Leu	tat Tyr	ggc Gly	1632
gag Glu 545	gat Asp	gtt Val	tac Tyr	gtt Val	ggc Gly 550	tac Tyr	aag Lys	ttc Phe	tac Tyr	gac Asp 555	aaa Lys	acg Thr	gag Glu	agg Arg	gag Glu 560	1680
gtt Val	ctg Leu	ttt Phe	cct Pro	ttt Phe 565	ggg Gly	cat His	ggc Gly	ctg Leu	tct Ser 570	tac Tyr	gct Ala	acc Thr	ttc Phe	aag Lys 575	ctc Leu	1728
cca Pro	gat Asp	tct Ser	acc Thr 580	gtg Val	agg Arg	acg Thr	gtc Val	ccc Pro 585	gaa Glu	acc Thr	ttc Phe	cac His	ccg Pro 590	gac Asp	cag Gln	1776
ccc Pro	aca Thr	gta Val 595	gcc Ala	att Ile	gtc Val	aag Lys	atc Ile 600	aag Lys	aac Asn	acg Thr	agc Ser	agt Ser 605	gtc Val	ccg Pro	ggc Gly	1824
gcc Ala	cag Gln 610	çtc Val	ctg Leu	cag Gln	tta Leu	tac Tyr 615	att Ile	tcg Ser	gcc Ala	cca Pro	aac Asn 620	tcg Ser	cct Pro	aca Thr	cat His	1872
cgc Arg 625	ccg Pro	gtc Val	aag Lys	gag Glu	ctg Leu 630	cac His	gga Gly	ttc Phe	gaa Glu	aag Lys 635	gtg Val	tat Tyr	ctt Leu	gaa Glu	gct Ala 640	1920
ggc Gly	gag Glu	gag Glu	aag Lys	gag Glu 645	gta Val	caa Gln	ata Ile	ccc Pro	att Ile 650	gac Asp	cag Gln	tac Tyr	gct Ala	act Thr 655	agc Ser	1968
t t c Phe	tgg Trp	gac Asp	gag Glu 660	att Ile	gag Glu	agc Ser	atg Met	tgg Trp 665	aag Lys	agc Ser	gag Glu	agg Arg	ggc Gly 670	att Ile	tat Tyr	2016
gat Asp	gtg Val	ctt Leu 675	gta Val	gga Gly	ttc Phe	t c g Se r	agt Ser 680	cag Gln	gaa Glu	atc Ile	tcg Ser	ggc Gly 685	aag Lys	ggg Gly	aag Lys	2064
ct q Le u	att Ile 690	gtg Val	cct Pro	gaa Glu	acg Thr	cga Arg 695	ttc Phe	tgg Trp	atg Met	ggg Gly	ctg Leu 700	tag				2103

- <210> 92 <211> 700 <212> PRT <213> Artificial Sequence <220> <223> Synthetic Construct <400> 92
- Phe Val Ala Asn Asp Gln Glu His Glu Arg Arg Ala Val Asp Cys Leu 20 25 30
- Ile Thr Gln Arg Ala Leu Arg Glu Val Tyr Leu Arg Pro Phe Gln Ile 35 40 45
- Val Ala Arg Asp Ala Arg Pro Gly Ala Leu Met Thr Ser Tyr Asn Lys 50 55 60
- Val Asn Gly Lys His Val Ala Asp Ser Ala Glu Phe Leu Gln Gly Ile 65 70 75 80
- Leu Arg Thr Glu Trp Asn Trp Asp Pro Leu Ile Val Ser Asp Trp Tyr 85 90 95
- Gly Thr Tyr Thr Thr Ile Asp Ala Ile Lys Ala Gly Leu Asp Leu Glu 100 105 110
- Met Pro Gly Val Ser Arg Tyr Arg Gly Lys Tyr Ile Glu Ser Ala Leu 115 120 125
- Gln Ala Arg Leu Leu Lys Gln Ser Thr Ile Asp Glu Arg Ala Arg Arg 130 135 140
- Val Leu Arg Phe Ala Gln Lys Ala Ser His Leu Lys Val Ser Glu Val
- Glu Gln Gly Arg Asp Phe Pro Glu Asp Arg Val Leu Asn Arg Gln Ile 165 170 175
- Cys Gly Ser Ser Ile Val Leu Leu Lys Asn Glu Asn Ser Ile Leu Pro 180 185 190

Leu	Pro	Lvs	Ser	Val	Lys	Lys	Val	Ala	Leu	Val	Gly	Ser	His	Val	Arg
		195			•	=	200					205			

- Leu Pro Ala Ile Ser Gly Gly Gly Ser Ala Ser Leu Val Pro Tyr Tyr 210 215 220
- Ala Ile Ser Leu Tyr Asp Ala Val Ser Glu Val Leu Ala Gly Ala Thr 225 230 235
- Ile Thr His Glu Val Gly Ala Tyr Ala His Gln Met Leu Pro Val Ile $\frac{245}{255}$
- Asp Ala Met Ile Ser Asn Ala Val Ile His Phe Tyr Asn Asp Pro Ile 260 265 270
- Asp Val Lys Asp Arg Lys Leu Leu Gly Ser Glu Asn Val Ser Ser Thr 275 280 285
- Ser Phe Gln Leu Met Asp Tyr Asn Asn Ile Pro Thr Leu Asn Lys Ala 290 300
- Met Phe Trp Gly Thr Leu Val Gly Glu Phe Ile Pro Thr Ala Thr Gly 305 310 315
- Ile Trp Glu Phe Gly Leu Ser Val Phe Gly Thr Ala Asp Leu Tyr Ile 325 330 335
- Asp Asn Glu Leu Val Ile Glu Asn Thr Thr His Gln Thr Arg Gly Thr 340 345 350
- Ala Phe Phe Gly Lys Gly Thr Thr Glu Lys Val Ala Thr Arg Arg Met 355 360 365
- Val Ala Gly Ser Thr Tyr Lys Leu Arg Leu Glu Phe Gly Ser Ala Asn 370 375 380
- Thr Thr Lys Met Glu Thr Thr Gly Val Val Asn Phe Gly Gly Gly Ala 385 390 395 400
- Val His Leu Gly Ala Cys Leu Lys Val Asp Pro Gln Glu Met Ile Ala 405 410 415
- Arg Ala Val Lys Ala Ala Ala Asp Ala Asp Tyr Thr Ile Ile Cys Thr 420 425 430

· Page 107

- Gly Leu Ser Gly Glu Trp Glu Ser Glu Gly Phe Asp Arg Pro His Met 435
- Asp Leu Pro Pro Gly Val Asp Thr Met Ile Ser Gln Val Leu Asp Ala 450 455 460
- Ala Pro Asn Ala Val Val Val Asn Gln Ser Gly Thr Pro Val Thr Met 465 470 470 480
- Ser Trp Ala His Lys Ala Lys Ala Ile Val Gln Ala Trp Tyr Gly Gly 485 490 495
- Asn Glu Thr Gly His Gly Ile Ser Asp Val Leu Phe Gly Asn Val Asn 500 505
- Pro Ser Gly Lys Leu Ser Leu Ser Trp Pro Val Asp Val Lys His Asn 515 520 525
- Pro Ala Tyr Leu Asn Tyr Ala Ser Val Gly Gly Arg Val Leu Tyr Gly 530 535 540
- Glu Asp Val Tyr Val Gly Tyr Lys Phe Tyr Asp Lys Thr Glu Arg Glu 545 550 555 560
- Val Leu Phe Pro Phe Gly His Gly Leu Ser Tyr Ala Thr Phe Lys Leu 565 570 575
- Pro Asp Ser Thr Val Arg Thr Val Pro Glu Thr Phe His Pro Asp Gln 580 585 590
- Pro Thr Val Ala Ile Val Lys Ile Lys Asn Thr Ser Ser Val Pro Gly 595 600 605
- Ala Gln Val Leu Gln Leu Tyr Ile Ser Ala Pro Asn Ser Pro Thr His 610 620
- Arg Pro Val Lys Glu Leu His Gly Phe Glu Lys Val Tyr Leu Glu Ala 625 630 635 640
- Gly Glu Glu Lys Glu Val Gln Ile Pro Ile Asp Gln Tyr Ala Thr Ser 645 650 655

Phe Trp Asp Glu Ile Glu Ser Met Trp Lys Ser Glu Arg Gly-Ile Tyr 660

Asp Val Leu Val Gly Phe Ser Ser Gln Glu Ile Ser Gly Lys Gly Lys 680

Leu Ile Val Pro Glu Thr Arg Phe Trp Met Gly Leu 695

<210> 93

<211> 1496 <212> DNA

<213> Artificial Sequence

<220>

<223> Maize optimized CBHI

<400> 93 tgcagtoogo otgcaccoto cagtoogaga cocaccogoo gotcacctgg cagaagtgot 60 120 ceteeggegg cacetgeace cageagaceg geteegtggt gategaegee aactggeget ggacccacgo caccaactoo tocaccaact gotacgacgg caacacetgg tootocacco 180 tetgecegga caacgagace tgegecaaga actgetgeet egaeggegee geetaegeet 240 300 ccacctacgg cgtgaccacc tccggcaact ccctctccat cggcttcgtg acccagtccg cccagaagaa cgtgggcgcc cgcctctacc tcatggcctc cgacaccacc taccaggagt 360 traccorrect eggeaacgag trefeetteg acgtggacgt greecagete ecgtgeggee 420 tcaacggcgc cetetactic gigtecatgg acgecgacgg eggegigtee aagtaceega 480 ccaacacege eggegeeaag taeggeaeeg getaetgega eteccagtge eegegegaee 540 tcaagttcat caacggccag gccaacgtgg agggctggga gccgtcctcc aacaacgcca 600 acaccageat eggeggeeae ggeteetget geteegagat ggacatetgg gaggeeaact 660 720 ccatctccga ggccctcacc ccgcacccgt gcaccaccgt gggccaggag atctgcgagg gegacgoetg eggeogeace tacteegaca accoetacgo eggeacetge gacceggacg 780 840 actocquetg gaaccogtac egecteggea acaceteett etaeggeeeg ggeteeteet tcaccetega caccaccaag aageteaceg tggtgaceca gttegagace teeggegeca 900 tcaaccgcta ctacgtgcag aacggcgtga ccttccagca gccgaacgcc gagctcggct 960 cctactccgg caacgagete aacgacgaet actgcaccge cgaggaggee gagtteggeg 1020 getectectt etecgacaag ggeggeetea eccagtteaa gaaggeeace teeggeggea 1080 tggtgctcgt gatgtccctc tgggacgact actacgccaa catgctctgg ctcgactcca 1140

Page 109

cctacccgac caa	egagace te	ctccaccc	cgggcgccgt	gcgcggctcc	tgctccacct	1200
ceteeggegt gee	ggcccag gt	ggagtccc	agtccccgaa	cgccaaggtg	accttctcca	1260
acatcaagtt cgg	cccgatc gg	getecaceg	gcaacccgtc	cggcggcaac	ccgccgggcg	1320
gcaacccgcc ggg	caccacc ac	caccogcc	gcccggccac	caccaccggc	tecteccegg	1380
gecegaceca gte	ccactac gg	gccagtgcg	gcggcatcgg	ctactccggc	ccgaccgtgt	1440
gegeeteegg cace	cacetge ca	iggtgctca	acccgtacta	ctcccagtgc	ctctag	1496
<220>	ial Sequen otimized C					
<400> 94 atggtgccgc tcg	angagco cc	aggeetge	tectecgtgt	ggggccagtg	cggcggccag	60
aactggtocg goo						120
tactoccagt gcc						180
acctocogog tgt						240
accaccacco gog						300
gtgggcgtga ccc						360
cogtocottoa cog						420
atgtggctcg aca						480
accgccaaca aga						540
cgcgactgcg ccg						600
aagtacaaga act						660
accetected to						720
aagtgcgcca acg						780
aacctocoga acg						840
gccaaccagg acc						900
cacaaccada acc						960
cogcogtect acad						1020
ggcccgctcc tcg						1080
ggcccgcccc reg	ccaacca cy	gurggree			9999-	

tccggcaagc	agccgaccgg	ccagcagcag	tggggcgact	ggtgcaacgt	gatcggcacc	1140
gacttcggca	teegeeegte	cgccaacacc	ggcgactccc	tcctcgactc	cttcgtgtgg	1200
gtgaagccgg	gcggcgagtg	cgacggcacc	tccgactcct	ccgccccgcg	cttcgactcc	1260
cactgcgccc	teceggaege	cctccagccg	gccccgcagg	ccggcgcctg	gttccaggcc	1320
tacttcgtgc	ageteeteae	caacgccaac	ccgtccttcc	tctag		1365
<220>	7 ificial Sequ ze optimizeo					
	cgggcacctc	caccccggag	gtgcacccga	agctcaccac	ctacaagtgc	60
accaagtccg	gcggctgcgt	ggcccaggac	acctccgtgg	tgctcgactg	gaactaccgc	120
tagatacaca	acgccaacta	caactcctgc	accgtgaacg	gcggcgtgaa	caccaccctc	180
tacccadaca	aggccacctg	cggcaagaac	tgcttcatcg	agggcgtgga	ctacgccgcc	240
tecogegtga	ccacctccgg	ctcctccctc	accatgaacc	agtacatgcc	gtcctcctcc	300
ągcągctact	cctccgtgtc	cccgcgcctc	tacctcctcg	actccgacgg	cgagtacgtg	360
atgctcaagc	tcaacggcca	ggagetetee	ttcgacgtgg	acctctccgc	cctcccgtgc	420
ggcgagaacg	geteceteta	cctctcccag	atggacgaga	acggcggcgc	caaccagtac	480
aacacegeeg	gogocaacta	cggctccggc	tactgcgacg	cccagtgccc	ggtgcagacc	540
tggcgcaacg	gcaccctcaa	cacctcccac	cagggcttct	gctgcaacga	gatggacatc	600
ctcgagggca	actcccgcgc	caacgccctc	accccgcact	cctgcaccgc	caccgcctgc	660
gacteegeeg	gatgaggatt	caacccgtac	ggctccggct	acaagtccta	ctacggcccg	720
gacgacaccg	tggacacctc	caagaccttc	accatcatca	cccagttcaa	caccgacaac	780
gactccccgt	ccggcaacct	cgtgtccatc	acccgcaagt	accagcagaa	cggcgtggac	840
atecegteeg	cccagccggg	cggcgacacc	atctcctcct	gcccgtccgc	ctccgcctac	900
ggcggcctcg	ccaccatggg	caaggccctc	tcctccggca	tggtgctcgt	gttctccatc	960
tggaacgaca	actcccagta	catgaactgg	ctcgactccg	gcaacgccgg	cccgtgctcc	1020
tccaccgagg	gcaacccgtc	caacaccctc	gccaacaacc	cgaacaccca	cgtggtgttc	1080
tccaacatcc	gctggggcga	catcggctcc	accaccaact	ccaccgcccc	gccgccgccg	1140

Page 111

coggettest ccaccacett etecaceace egeogeteet ecaccacete etect	ccccg 1200
tectigeacce agacecacty gagecagtge ageggeateg getacteegg etgea	agacc 1260
tgcacctccg gcaccacctg ccagtactcc aacgactact actcccagtg cctct	ag 1317
<210> 96 <211> 1401 <212> DNA <213> Artificial Sequence	
<220> <223> Maize optimized BGLII	•
<400> 96 atgeteeega aggaetteea gtggggette gecaeegeeg eetaceagat egagg	ggcgcc 60
gregaccage aeggeogegg ecegrecate rgggacacer rergegeeca geoge	ggcaag 120
ategeogacg geteeteegg egtgaeegee tgegaeteet acaacegeae egeeg	gaggac 180
atogocotoc toaagtooot oggogocaag tootacogot totocatoto otggt	cccgc 240
atcatecegg agggeggeeg eggegaegee gtgaaceagg eeggeatega ceaet	acgtg 300
aagttegtgg acgaeeteet egaegeegge atcaceeegt teateaceet ettee	actgg 360
gacetecegg agggeeteca ecagegetae ggeggeetee teaacegeae egagt	tcccg 420
ctcgacttcg agaactacge cegegtgatg ttccgegeee teeegaaggt gegea	actgg 480
atcaccttca acgageoget etgeteegee atceeggget acggeteegg cacet	tegee 540
ccgggccgcc agtccacctc cgagccgtgg accgtgggcc acaacatcct cgtgg	jcccac 600
ggccgcgccg tgaaggccta ccgcgacgac ttcaagccgg cctccggcga cggc	agatc 660
ggcatcgtgc tcaacggcga cttcacctac ccgtgggacg ccgccgaccc ggccg	gacaag 720
gaggeegeeg agegeegeet egagttette acegeetggt tegeegaeee gatet	acctc 780
ggcgactace eggeeteeat gegeaageag eteggegaee geeteeegae ettea	accccg 840
gaggagegeg cectegtgea eggetecaae gaettetaeg geatgaacea etaea	ectcc 900
aactacatee gecacegete etecceggee teegeegacg acacegtggg caace	gtggac 960
gtgctcttca ccaacaagca gggcaactgc atcggcccgg agacccagtc cccgt	ggctc 1020
egecegugeg cegeeggett eegegaette etegtgtgga tetecaageg etace	gctac 1080
ccgccgatct acgtgaccga gaacggcace tccatcaagg gcgagtccga cctcc	cgaag 1140
gagaagatee tegaggaega etteegegtg aagtaetaea acgagtaeat eegeg	ccatg 1200
graccacca ragagereda eddedtase dragadder actredecta drece	tcatg 1260

gacaacttog agtgggooga oggotaogtg accogottog gogtgacota ogtggactao	1320
gagaacggcc agaagcgctt cccgaagaag tccgccaagt ccctcaagcc gctcttcgac	1380
gageteateg eegeegeeta g	1401
<210> 97 <211> 2103 <212> DNA <213> Artificial Sequence	·
<220> <223> Maize optimized CEL3D	
<400> 97 atgatecteg getgegagte caceggegtg ateteegeeg tgaageaett egtggecaae	60
gaccaggage acgagegeeg egeegtggae tgeeteatea eccagegege ecteegegag	120
giglacetee geoegiteea galegiggee egegaegeee geoegggege eetcalgace	180
tectacaaca aggigaacgg caageacgig geegacteeg eegagiteet eeagggeate	240
ctccgcaccg agtggaactg ggacccgctc atcgtgtccg actggtacgg cacctacacc	300
accategacg ccateaagge eggeetegae etegagatge egggegtgte eegetacege	360
ggcaagtaca togagtoogo ootocaggoo ogootootoa agcagtooac catogacgag	420
egegeeegee gegtgeteeg ettegeeeag aaggeeteee aceteaaggt gteegaggtg	480
gagcagggee gegaetteee ggaggaeege gtgeteaace gecagatetg eggeteetee	540
atogogotoc toaagaacga gaactocato otocogotoo ogaagtoogt gaagaaggtg	600
geoctegigg geteceacgt gegeeteecg gecateteeg geggeggete egecteecte	660
gtgccgtact acgccatete cetetacgae geegtgteeg aggtgetege eggegeeace	720
atcacccacg aggtgggcgc ctacgcccac cagatgctcc cggtgatcga cgccatgatc	780
tecaaegeeg tgatecaett etacaaegae eegategaeg tgaaggaeeg caageteete	840
ggctccgaga acgtgtcctc cacctccttc cagctcatgg actacaacaa catcccgacc	900
ctcaacaagg ccatgttctg gggcaccctc gtgggcgagt tcatcccgac cgccaccggc	960
atotogogagi togocototo ogiqitogoo accocogaco totacatoga caacqagoto	1020
gtgatcgaga acaccaccca ccagacccgc ggcaccgcct tetteggcaa gggcaccacc	1080
gagaaggtgg ccacccgccg catggtggcc ggctccacct acaagctccg cctcgagttc	1140
ggctccgcca acaccaccaa gatggagacc accggcgtgg tgaacttcgg cggcggcgcc	1200
grycaccieg gegeetgeer caaggrygae eegeaggaga rgaregeeeg egeegtgaag	1260

gccgccgccg	acgccgacta	caccatcatc	tgcaccggcc	tctccggcga	gtgggagtcc	1320
gagggetteg	accgcccgca	catggacctc	ccgccgggcg	tggacaccat	gatctcccag	1380
gtgctcgacg	ccgccccgaa	cgccgtggtg	gtgaaccagt	ccggcacccc	ggtgaccatg	1440
tcctgggccc	acaaggccaa	ggccatcgtg	caggcctggt	acggcggcaa	cgagaccggc	1500
cacogcatct	ccgacgtgct	cttcggcaac	gtgaacccgt	ccggcaagct	ctccctctcc	1560
tggccggtgg	acgtgaagca	caacccggcc	tacctcaact	acgcctccgt	gggcggccgc	1620
gtgctctacg	gcgaggacgt	gtacgtgggc	tacaagttct	acgacaagac	cgagcgcgag	1680
gtactettee	cgttcggcca	cggcctctcc	tacgccacct	tcaagctccc	ggactccacc	1740
gtgcgcaccg	tgccggagac	cttccacccg	gaccageega	ccgtggccat	cgtgaagatc	1800
aagaacacct	cctccgtgcc	gggcgcccag	gtgctccagc	tctacatctc	cgccccgaac	1860
teccegaces	accgcccggt	gaaggagctc	cacggcttcg	agaaggtgta	cctcgaggcc	1920
ggcgaggaga	aggaggtgca	gatcccgatc	gaccagtacg	ccacctcctt	ctgggacgag	1980
atcgagtcca	tgtggaagtc	cgagcgcggc	atctacgacg	tgctcgtggg	cttctcctcc	2040
caggagatet	ccggcaaggg	caagctcatc	gtgccggaga	cccgcttctg	gatgggcctc	2100
tag						2103
<220>	mays rotein promo	oter				
<400> 98	attacttggg	agcaatggta	tgcaaatcct	ttgcatgtac	gcaaaactag	60
	caagttgtat					120
	caatattaca					180
	agtgaagctt					240
	aaaaatccac					300
	aaatgtttat					360
	ccaacaaagc					420
<210> 99 <211> 1189						

Page 114

```
<212> DNA
<213> artificial sequence
<223> synthetic ferulic acid esterase
atggccgcct ccctcccgac catgccgccg tccggctacg accaggtgcg caacggcgtg
                                                                       60
ccgcgcggcc aggtggtgaa catctcctac ttctccaccg ccaccaactc cacccgcccg
                                                                      120
gcccgcgtgt acctcccgcc gggctactcc aaggacaaga agtactccgt gctctacctc
                                                                      180
ctccacggca tcggcggctc cgagaacgac tggttcgagg gcggcggccg cgccaacgtg
                                                                      240
atogoogada acotoatogo ogagggdaag atcaagoogo toatoatogt gacoocgaac
                                                                      300
accaacgeeg eeggeeeggg categeegae ggetaegaga actteaceaa ggaceteete
                                                                      360
aactocotca tooogtacat ogagtocaac tactoogtgt acaccgacog ogagcacogo
                                                                      420
                                                                      480
gccatcgccg gcctctctat gggcggcggc cagtccttca acatcggcct caccaacctc
gacaagttog octabatogg ecogatotoe geogeocoga acacetacoe gaacgagege
                                                                      540
ctcttcccgg acggcggcaa ggccgcccgc gagaagetca agetcetett categeetge
                                                                      600
                                                                      660
ggcaccaacg actocotoat oggottoggo cagogogtgo acgagtactg ogtggocaac
                                                                      720
aacatcaacc acgtgtactg gctcatccag ggcggcggcc acgacttcaa cgtgtggaag
                                                                      780
cogggeotet ggaactteet coagatggee gacgaggeeg geeteacceg egacggeaac
acconggige egaceoegie ecogaageog gecaacacee geategagge egaggactae
                                                                      840
gacggcatca actoctocto catogagato atoggogtgo ogcoggaggg oggoogogo
                                                                      900
atoggotaca toacotoogg egactacete gtgtacaagt ecategaett eggeaacgge
                                                                      960
gecaceteet teaaggecaa ggtggecaae gecaacacet ceaacatega gettegeete
                                                                     1020
                                                                     1080
aacggcccga acggcaccct catcggcacc ctctccgtga agtccaccgg cgactggaac
acctacgagg ageagacetg etecatetee aaggtgaceg geateaacga cetetacete
                                                                     1140
                                                                     1188
gtgttcaagg gcccggtgaa catcgactgg ttcaccttcg gcgtgtag
<210> 100
<211> 395
<212> PRT
<213> artificial sequence
<223> synthetic ferulic acid esterase
```

<400> 100

Met Ala Ala Ser Leu Pro Thr Met Pro Pro Ser Gly Tyr Asp Gln Val 1 5 10

Arg Asn Gly Val Pro Arg Gly Gln Val Val Asn Ile Ser Tyr Phe Ser 20 25 30

Thr Ala Thr Asn Ser Thr Arg Pro Ala Arg Val Tyr Leu Pro Pro Gly 35 40 45

Tyr Ser Lys Asp Lys Lys Tyr Ser Val Leu Tyr Leu Leu His Gly Ile 50 55 60

Gly Gly Ser Glu Asn Asp Trp Phe Glu Gly Gly Gly Arg Ala Asn Val $_{65}$ 70 75 80

Ile Ala Asp Asn Leu Ile Ala Glu Gly Lys Ile Lys Pro Leu Ile Ile 85 90 95

Val Thr Pro Asn Thr Asn Ala Ala Gly Pro Gly Ile Ala Asp Gly Tyr 100 105 110

Glu Asn Phe Thr Lys Asp Leu Leu Asn Ser Leu Ile Pro Tyr Ile Glu 115 120 125

Ser Asn Tyr Ser Val Tyr Thr Asp Arg Glu His Arg Ala Ile Ala Gly 130 135 140

Leu Ser Met Gly Gly Gly Gln Ser Phe Asn Ile Gly Leu Thr Asn Leu 145 150 155 160

Asp Lys Phe Ala Tyr Ile Gly Pro Ile Ser Ala Ala Pro Asn Thr Tyr 165 170 175

Pro Asn Glu Arg Leu Phe Pro Asp Gly Gly Lys Ala Ala Arg Glu Lys 180 185 190

Leu Lys Leu Eu Fhe Ile Ala Cys Gly Thr Asn Asp Ser Leu Ile Gly 195 200 205

Phe Gly Gln Arg Val His Glu Tyr Cys Val Ala Asn Asn Ile Asn His 210 220

Val Tyr Trp Leu Ile Gln Gly Gly Gly His Asp Phe Asn Val Trp Lys 225 230 235 240

Page 116

Pro	Gly	Leu	Trp	Asn 245	Phe	Leu	Gln	Met	Ala 250	Asp	Glu	Ala	Gly	Leu 255	Thr		
Arg	Asp	Gly	Asn 260	Thr	Pro	Val	Pro	Thr 265	Pro	Ser	Pro	Lys	Pro 270	Ala	Asn		
Thr	Arg	Ile 275	Glu	Ala	Glu	Asp	Tyr 280	Asp	Gly	Ile	Asn	Ser 285	Ser	Ser	Ile		
Glu	Ile 290	Ile	Gly	Val	Pro	Pro 295	Glu	Gly	Gly	Arg	Gly 300	Ile	Gly	Tyr	Ile		
Thr 305	Ser	Gly	Asp	Tyr	Leu 310	Val	Tyr	Lys	Ser	11e 315	Asp	Phe	Gly	Asn	Gly 320		
Ala	Thr	Ser	Phe	Lys 325	Ala	Lys	Val	Ala	Asn 330	Ala	Asn	Thr	Ser	Asn 335	Ile		
Glu	Leu	Arg	Leu 340	Asn	Gly	Pro	Asn	Gly 345	Thr	Leu	Ile	Gly	Thr 350	Leu	Ser		
Val	Lys	Ser 355	Thr	Gly	Asp	Trp	Asn 360	Thr	Tyr	Glu	Glu	Gln 365	Thr	Cys	Ser		
Ile	Ser 370	Lys	Val	Thr	Gly	Ile 375	Asn	Asp	Leu	Tyr	Leu 380	Val	Phe	Lys	Gly		
Pro 385	Val	Asn	Ile	Asp	Trp 390	Phe	Thr	P'ne	G1 y	Val 395							
<210 <211 <212 <213	> 	101 1188 DMA arti!	ficia	el se	equer	nce		÷									
<220 <221		olasm	nıd l	3036	5												
<400 atgc		101 ect 6	cctc	ccga	ec ca	atgco	gee	g tco	ggct	acg	acca	aggto	gcg (caacq	gcgtg		60
ccad	gcgo	gec a	ggt	ggtga	a ca	tcto	ctac	tto	tcca	ccg	ccac	ccaac	ctc d	cacco	gcccg	1	20
gcco	gcgt	gt a	eccto	ccgc	c gç	ggcta	ectco	aaq	gaca	aga	agta	ctco	egt q	gctct	acctc	1	80
ctco	acgo	gca t	cggc	gget	c cç	jagaa	acgad	t g	ttc	jagg	gcg	gegge	ccg (gcca	acgtg	2	240

- (gaca	acctcatcgc	cgagggcaag	atcaagccgc	tcatcatcgt	gaccccgaac	300
20	geeg	ccggcccggg	catogoogac	ggctacgaga	acttcaccaa	ggacctcctc	360
20	ctca	tecegtacat	cgagtccaac	tactccgtgt	acaccgaccg	cgagcaccgc	420
= 0	gccg	gcctctctat	gggcggcggc	cagteettea	acatcggcct	caccaacctc	480
j t	tcg	cctacatcgg	cccgatctcc	gccgccccga	acacctaccc	gaacgagcgc	540
- (cgg	acggcggcaa	ggccgcccgc	gagaagctca	agctcctctt	catcgcctgc	600
2 8	aacg	actccctcat	cggcttcggc	cagcgcgtgc	acgagtactg	cgtggccaac	660
e a	sacc	acgtgtactg	gctcatccag	ggcggcggcc	acgacttcaa	cgtgtggaag	720
2 (ctct	ggaacttcct	ccagatggcc	gacgaggccg	gcctcacccg	cgacggcaac	780
90	gtgc	cgaccccgtc	cccgaagccg	gccaacaccc	gcatcgaggc	cgaggactac	840
cá	atca	actectecte	catcgagatc	atcggcgtgc	cgccggaggg	cggccgcggc	900
= 1	caca	tcacctccgg	cgactacctc	gtgtacaagt	ccatcgactt	cggcaacggc	960
21	cct	tcaaggccaa	ggtggccaac	gccaacacct	ccaacatcga	gcttcgcctc	1020
20	ccga	acggcaccct	catcggcacc	ctctccgtga	agtccaccgg	cgactggaac	1080
20	gagg	agcagacctg	ctccatctcc	aaggtgaccg	gcatcaacga	cctctacctc	1140
: a	aagg	gcccggtgaa	catcgactgg	ttcaccttcg	gcgtgtag		1188

<210> 102

<211> 395 <212> PRT <213> artificial sequence

<220>

<223> plasmid 13036

<400> 102

Met Ala Ala Ser Leu Pro Thr Met Pro Pro Ser Gly Tyr Asp Gln Val 1 5 15

Arg Asn Gly Val Pro Arg Gly Gln Val Val Asn Ile Ser Tyr Phe Ser 20 25 30

Thr Ala Thr Asn Ser Thr Arg Pro Ala Arg Val Tyr Leu Pro Pro Gly

Tyr Ser Lys Asp Lys Lys Tyr Ser Val Leu Tyr Leu Leu His Gly Ile 50 55

Page 118

- Gly Gly Ser Glu Asn Asp Trp Phe Glu Gly Gly Gly Arg Ala Asn Val 70 75 80
- Ile Ala Asp Asn Leu Ile Ala Glu Gly Lys Ile Lys Pro Leu Ile Ile 85 90 95
- Val Thr Pro Asn Thr Asn Ala Ala Gly Pro Gly Ile Ala Asp Gly Tyr $100 \,$
- Giu Asn Phe Thr Lys Asp Leu Leu Asn Ser Leu Ile Pro Tyr Ile Glu 115 120 125
- Ser Asn Tyr Ser Val Tyr Thr Asp Arg Glu His Arg Ala Ile Ala Gly 130 140
- Leu Ser Met Gly Gly Gly Gln Ser Phe Asn Ile Gly Leu Thr Asn Leu 145 150 160
- Asp Lys Phe Ala Tyr Ile Gly Pro Ile Ser Ala Ala Pro Asn Thr Tyr 165 170 175
- Pro Asn Glu Arg Leu Phe Pro Asp Gly Gly Lys Ala Ala Arg Glu Lys 180 185 190
- Leu Lys Leu Phe Ile Ala Cys Gly Thr Asn Asp Ser Leu Ile Gly 195 200 205
- Phe Gly Gln Arg Val His Glu Tyr Cys Val Ala Asn Asn Ile Asn His 210 215 220
- Val Tyr Trp Leu Ile Gln Gly Gly Gly His Asp Phe Asn Val Trp Lys 225 230 235
- Pro Gly Leu Trp Asn Phe Leu Gln Met Ala Asp Glu Ala Gly Leu Thr 245 250 255
- Arg Asp Gly Asn Thr Pro Val Pro Thr Pro Ser Pro Lys Pro Ala Asn 260 265 270
- Thr Arg Ile Glu Ala Glu Asp Tyr Asp Gly Ile Asn Ser Ser Ser Ile 275 280 285

Glu Ile Ile Gly Val Pro Pro Glu Gly Gly Arg Gly Ile Gly Tyr Ile Thr Ser Gly Asp Tyr Leu Val Tyr Lys Ser Ile Asp Phe Gly Asn Gly 315 310 Ala Thr Ser Phe Lys Ala Lys Val Ala Asn Ala Asn Thr Ser Asn Ile 330 335 325 Glu Leu Arg Leu Asn Gly Pro Asn Gly Thr Leu Ile Gly Thr Leu Ser Val Lys Ser Thr Gly Asp Trp Asn Thr Tyr Glu Glu Gln Thr Cys Ser 360 Ile Ser Lys Val Thr Gly Ile Asn Asp Leu Tyr Leu Val Phe Lys Gly 375 380 Pro Val Asn Ile Asp Trp Phe Thr Phe Gly Val 395 390 <210> 103 <211> 1245 <212> DNA <213> artificial sequence <220> <223> plasmid 13038 <400> 103 atgagggtgt tgctcgttgc cctcgctctc ctggctctcg ctgcgagcgc cacctccatg 60 geogeotece tecogaceat geogeogtee ggetacgace aggtgegeaa eggegtgeeg 120 cgcggccagg tggtgaacat ctcctacttc tccaccgcca ccaactccac ccgcccggcc 180 240 egegigtace tecegeeggg elacteeaag gacaagaagt acteegiget elaceteete cacqqcatcq gcqqctccqa gaacqactqq ttcqaqqqcq qcqqccqcqc caacqtqatc 300 googacaaco toatogooga gggcaagato aagoogotoa toatogtgac coogaacaco 360 aacgccgccg gcccgggcat cgccgacggc tacgagaact tcaccaagga cctcctcaac 420 teceteatee egtacatega gtecaactae teegtgtaca eegacegega geacegegee 480 atcgccggcc tototatggg cggcggccag toottcaaca tcggcctcac caacctcgac 540 aagttegeet acateggeee gateteegee geeeegaaca eetaeeegaa egagegeete 600 660 ttcccggacg gcggcaaggc cgcccgcgag aagetcaage teetetteat cgcctgcggc

accaacgact	ccctcatcgg	cttcggccag	cgcgtgcacg	agtactgcgt	ggccaacaac	720
atcaaccacg	tgtactggct	catccagggc	ggcggccacg	acttcaacgt	gtggaagccg	780
ggcctctgga	acttcctcca	gatggccgac	gaggccggcc	tcacccgcga	cggcaacacc	840
ccggtgccga	ccccgtcccc	gaagccggcc	aacacccgca	tcgaggccga	ggactacgac	900
ggcatcaact	cctcctccat	cgagatcatc	ggcgtgccgc	cggagggcgg	ccgcggcatc	960
ggctacatca	cctccggcga	ctacctcgtg	tacaagtcca	tcgacttcgg	caacggcgcc	1020
acctccttca	aggccaaggt	ggccaacgcc	aacacctcca	acatcgagct	tcgcctcaac	1080
ggcccgaacg	gcaccctcat	cggcaccctc	tccgtgaagt	ccaccggcga	ctggaacacc	1140
tacgaggage	agacctgctc	catctccaag	gtgaccggca	tcaacgacct	ctacctcgtg	1200
ttcaagggcc	cçgtgaacat	cgactggttc	accttcggcg	tgtag	·	1245

<210> 104 <211> 414 <212> PRT <213> artificial sequence

<220>

<223> plasmid 13038 aa

Met Arg Val Leu Leu Val Ala Leu Ala Leu Leu Ala Leu Ala Ala Ser

Ala Thr Ser Met Ala Ala Ser Leu Pro Thr Met Pro Pro Ser Gly Tyr

Asp Gln Val Arg Asn Gly Val Pro Arg Gly Gln Val Val Asn Ile Ser 40

Tyr Phe Ser Thr Ala Thr Asn Ser Thr Arg Pro Ala Arg Val Tyr Leu

Pro Pro Gly Tyr Ser Lys Asp Lys Lys Tyr Ser Val Leu Tyr Leu Leu 65 70 80

His Gly Ile Gly Gly Ser Glu Asn Asp Trp Phe Glu Gly Gly Arg 85 90 95

Ala Asn Val Ile Ala Asp Asn Leu Ile Ala Glu Gly Lys Ile Lys Pro

Page 121

Leu	lle	11e 115	Val	Thr	Pro	Asn	Thr 120	Asn	Ala	Ala	Gly	Pro 125	Gly	Ile	Ala
Asp	Gly 130	Tyr	Glu	Asn	Phe	Thr 135	Lys	Asp	Leu	Leu		Ser	Leu	Ile	Pro

- Tyr Ile Glu Ser Asn Tyr Ser Val Tyr Thr Asp Arg Glu His Arg Ala 145 150 155 160
- Ile Ala Gly Leu Ser Met Gly Gly Gly Gln Ser Phe Asn Ile Gly Leu
- Thr Asn Leu Asp Lys Phe Ala Tyr Ile Gly Pro Ile Ser Ala Ala Pro 180 185 190
- Asn Thr Tyr Pro Asn Glu Arg Leu Phe Pro Asp Gly Gly Lys Ala Ala 195 200 205
- Arg Glu Lys Leu Lys Leu Leu Phe Ile Ala Cys Gly Thr Asn Asp Ser 210 215 220
- Leu Ile Gly Phe Gly Gln Arg Val His Glu Tyr Cys Val Ala Asn Asn 225 230 235
- Ile Asn His Val Tyr Trp Leu Ile Gln Gly Gly Gly His Asp Phe Asn 245 250 255
- Vai Trp Lys Pro Gly Leu Trp Asn Phe Leu Gln Met Ala Asp Glu Ala 260 265 270
- Gly Leu Thr Arg Asp Gly Asn Thr Pro Val Pro Thr Pro Ser Pro Lys 275 280 285
- Pro Ala Asn Thr Arg Ile Glu Ala Glu Asp Tyr Asp Gly Ile Asn Ser 290 295 300
- Ser Ser Ile Glu Ile Ile Gly Val Pro Pro Glu Gly Gly Arg Gly Ile 305 310 315 320
- Gly Tyr Ile Thr Ser Gly Asp Tyr Leu Val Tyr Lys Ser Ile Asp Phe 325 335

Gly Asn Gly Ala Thr Ser Phe Lys Ala Lys Val Ala Asn Ala Asn Thr 340 345 350

Ser Asn Ile Glu Leu Arg Leu Asn Gly Pro Asn Gly Thr Leu Ile Gly 355 360 365

Thr Leu Ser Val Lys Ser Thr Gly Asp Trp Asn Thr Tyr Glu Glu Gln 370 380

Thr Cys Ser Ile Ser Lys Val Thr Gly Ile Asn Asp Leu Tyr Leu Val 385 390 395 400

Phe Lys Gly Pro Val Asn Ile Asp Trp Phe Thr Phe Gly Val 405

<210> 105 <211> 1425

<212> DNA

<213> artificial sequence

<220>

<223> plasmid 13039

<400> 105 atgctggcgg ctctggccac gtcgcagctc gtcgcaacgc gcgccggcct gggcgtcccg 60 gacqcqtcca cqttccqccq cqqcqccqcq caqqqcctqa qqqqqcccq qqcqtcqqcq 120 geggeggaca egeteageat geggaceage gegegegegg egeceaggea ecageaceag 180 caggogogoc goggggocag gttocogtog ctogtogtgt gogcoagogo oggogocatg 240 googootooc tooogaccat googoogtoo ggotacgaco aggtgogcaa oggogtgoog 300 egeggecagg tggtgaacat etectaette tecacegeca ceaactecae eegeceggee 360 cgcgtgtacc tcccgccggg ctactccaag gacaagaagt actccgtgct ctacctcctc 420 cacagocated geogeteega gaacgaeteg ttegagggeg geggeegege caacgtgate 480 geogacaaco teategooga gggeaagate aageogetea teategtgae eeegaacace 540 aacgccgccg gcccgggcat cgccgacggc tacgagaact tcaccaagga cctcctcaac 600 teceteatee egtacatega gtecaactae teegtgtaca eegacegega geacegegee 660 arcgccggcc tetetarggg eggeggecag teetteaaca teggeeteae caacetegae 720 aagttogoot acatoggood gatotoogoo goodogaaca cotacoogaa ogagogooto 780 ttcccggacg gcggcaaggc cgcccgcgag aagctcaagc tcctcttcat cgcctgcggc 840 accaacgact contrategy officeacy egegtiques agractigegt ggccaacaac 900

atcaaccacg	tgtactggct	caticcagggc	ggcggccacg	acttcaacgt	gtggaagccg	960
ggcctctgga	acttcctcca	gatggccgac	gaggccggcc	tcacccgcga	cggcaacacc	1020
ccggtgccga	ccccgtcccc	gaagccggcc	aacacccgca	tcgaggccga	ggactacgac	1080
ggcatcaact	cctcctccat	cgagatcatc	ggcgtgccgc	cggagggcgg	ccgcggcatc	1140
ggctacatca	cctccggcga	ctacctcgtg	tacaagtcca	tcgacttcgg	caacggcgcc	1200
acctccttca	aggccaaggt	ggccaacgcc	aacacctcca	acatcgagct	tcgcctcaac	1260
ggcccgaacg	gcaccctcat	cggcaccctc	tccgtgaagt	ccaccggcga	ctggaaçacc	1320
tacgaggagc	agacctgctc	catctccaag	gtgaccggca	tcaacgacct	ctacctcgtg	1380
ttcaagggcc	cogtgaacat	cgactggttc	accttcggcg	tgtag		1425

<210> 106 <211> 474 <212> PRT <213> artificial sequence

<220>

<223> plasmid 13039 aa

<400> 106

Met Leu Ala Ala Leu Ala Thr Ser Gln Leu Val Ala Thr Arg Ala Gly

Leu Gly Val Pro Asp Ala Ser Thr Phe Arg Arg Gly Ala Ala Gln Gly

Leu Arg Gly Ala Arg Ala Ser Ala Ala Ala Asp Thr Leu Ser Met Arg

Thr Ser Ala Arg Ala Ala Pro Arg His Gln His Gln Gln Ala Arg Arg

Gly Ala Arg Phe Pro Ser Leu Val Val Cys Ala Ser Ala Gly Ala Met

Ala Ala Ser Leu Pro Thr Met Pro Pro Ser Gly Tyr Asp Gln Val Arg

Asn Gly Val Pro Arg Gly Gln Val Val Asn Ile Ser Tyr Phe Ser Thr 105

Ala Thr Asn Ser Thr Arg Pro Ala Arg Val Tyr Leu Pro Pro-Gly Tyr 115 120 125

Ser Lys Asp Lys Lys Tyr Ser Val Leu Tyr Leu Leu His Gly Ile Gly 130 135 140

Gly Ser Glu Asn Asp Trp Phe Glu Gly Gly Gly Arg Ala Asn Val Ile 145 150 160

Ala Asp Asn Leu Ile Ala Glu Gly Lys Ile Lys Pro Leu Ile Ile Val

Thr Pro Asn Thr Asn Ala Ala Gly Pro Gly Ile Ala Asp Gly Tyr Glu 180 185 190

Asn Phe Thr Lys Asp Leu Leu Asn Ser Leu Ile Pro Tyr Ile Glu Ser 195 200 205

Asn Tyr Ser Val Tyr Thr Asp Arg Glu His Arg Ala Ile Ala Gly Leu 210 215 220

Ser Met Gly Gly Gly Gln Ser Phe Asn Ile Gly Leu Thr Asn Leu Asp 225 230 235 240

Lys Phe Ala Tyr Ile Gly Pro Ile Ser Ala Ala Pro Asn Thr Tyr Pro 245 255

Asn Glu Arg Leu Phe Pro Asp Gly Gly Lys Ala Ala Arg Glu Lys Leu 260 265 270

Lys Leu Phe Ile Ala Cys Gly Thr Asn Asp Ser Leu Ile Gly Phe 275 280 285

Gly Gln Arg Val His Glu Tyr Cys Val Ala Asn Asn Ile Asn His Val 290 295 300

Tyr Trp Leu Ile Gln Gly Gly Gly His Asp Phe Asn Val Trp Lys Pro 305 310 315

Gly Leu Trp Asn Phe Leu Gln Met Ala Asp Glu Ala Gly Leu Thr Arg 325 330 335

Page 125

Arg	Ile	Glu 355	Ala	Glu	Asp	Tyr	Asp 360	Gly	Ile	Asn	Ser	Ser 365	Ser	Ile	Glu	
Ile	Ile 370	Gly	Val	Pro	Pro	Glu 375	Gly	Gly	Arg	Gly	Ile 380	Gly	Tyr	Ile	Thr	
Ser 385	Gly	Asp	Tyr	Leu	Val 390	Tyr	Lys	Ser	Ile	Asp 395	Phe	Gly	Asn	Gly	Ala 400	
Thr	Ser	Phe	Lys	Ala 405	Lys	Val	Ala	Asn	Ala 410	Asn	Thr	Ser	Asn	Ile 415	Glu	
Leu	Arg	Leu	Asn 420	Gly	Pro	Asn	Gly	Thr 425	Leu	Ile	Gly	Thr	Leu 430	Ser	Val	
Lys	Ser	Thr 435	Gly	Asp	Trp	Asn	Thr 440	Tyr	Glu	Glu	Gln	Thr 445	Cys	Ser	Ile	
Ser	Lys 450	Val	Thr	Gly	Ile	Asn 455	Asp	Leu	Tyr	Leu	Val 460	Phe	Lys	Gly	Pro	
Val 465	Asn	Ile	Asp	Trp	Phe 470	Thr	Phe	Gly	Val							
<210 <210 <210 <210	1> : ?> !	107 1263 DNA arti:	ficia	el se	equer	nce										
<22(<22)		olası	mid :	1334	7											
<400 atg	39991 3>	107 Egt 1	tgct	egtte	ge e	tcg	ctct	c cto	ggeto	ctcg	ctg	cgago	ege (cacc	tccatg	60
gcc	geet	ccc 1	ccc	gacca	at go	cgc	cgtc	c gg	ctaco	gacc	agg	tgcg	caa (cggc	gtgccg	120
cgc	ggcca	agg 1	tggte	gaaca	at c	tocta	ectto	c tc	cacco	gcca	ccaa	actc	cac (ccgc	ccggcc	180
cgc	gtgta	acc 1	tccc	geege	gg ct	tacto	ccaa	g gad	caaga	agt	acto	ccgt	gct	ctac	ctcctc	240
cace	ggcai	tcg (gcgg	ctcc	ga ga	aacga	actgo	g tto	gage	ggcg	gcg	geege	cgc (caac	gtgatc	300
															aacacc	360
															ctcaac	420
															cgcgcc	480

ategeeggee	tctctatggg	cggcggccag	tccttcaaca	tcggcctcac	caacctcgac	540
aagttcgcct	acatcggccc	gatctccgcc	gccccgaaca	cctacccgaa	cgagcgcctc	600
ttcccggacg	gcggcaaggc	cgcccgcgag	aagctcaagc	tcctcttcat	cgcctgcggc	660
accaacgact	ccctcatcgg	cttcggccag	cgcgtgcacg	agtactgcgt	ggccaacaac	720
atcaaccacg	tgtactggct	catccagggc	ggcggccacg	acttcaacgt	gtggaagccg	780
ggcctctgga	acttcctcca	gatggccgac	gaggccggcc	tcacccgcga	cggcaacacc	840
ccggtgccga	ccccgtcccc	gaagccggcc	aacacccgca	tcgaggccga	ggactacgac	900
ggcatcaact	cctcctccat	cgagatcatc	ggcgtgccgc	cggagggcgg	ccgcggcatc	960
ggctacatca	cctccggcga	ctacctcgtg	tacaagtcca	tcgacttcgg	caacggcgcc	1020
acctccttca	aggccaaggt	ggccaacgcc	aacacctcca	acatcgagct	tcgcctcaac	1080
ggcccgaacg	gcaccctcat	cggcaccctc	tccgtgaagt	ccaccggcga	ctggaacacc	1140
tacqaggagc	agacctgctc	catctccaag	gtgaccggca	tcaacgacct	ctacctcgtg	1200
ttcaagggce	cggtgaacat	cgactggttc	accttcggcg	tgtccgagaa	ggacgaactc	1260
tag						1263

<210> 108 <211> 420 <212> PRT <213> artificial sequence

<223> plasmid 13347

<400> 108

Met Arg Val Leu Leu Val Ala Leu Ala Leu Leu Ala Leu Ala Ala Ser

Ala Thr Ser Met Ala Ala Ser Leu Pro Thr Met Pro Pro Ser Gly Tyr

Asp Gln Val Arg Asn Gly Val Pro Arg Gly Gln Val Val Asn Ile Ser

Tyr Phe Ser Thr Ala Thr Asn Ser Thr Arg Pro Ala Arg Val Tyr Leu

Pro Pro Gly Tyr Ser Lys Asp Lys Lys Tyr Ser Val Leu Tyr Leu Leu 70

His Gly Ile Gly Gly Ser Glu Asn Asp Trp Phe Glu Gly Gly Arg 85 90

Ala Asn Val Ile Ala Asp Asn Leu Ile Ala Glu Gly Lys Ile Lys Pro 100 105 110

Leu Ile Ile Val Thr Pro Asn Thr Asn Ala Ala Gly Pro Gly Ile Ala 115 120 125

Asp Gly Tyr Glu Asn Phe Thr Lys Asp Leu Leu Asn Ser Leu Ile Pro 130 135 140

Tyr Ile Glu Ser Asn Tyr Ser Val Tyr Thr Asp Arg Glu His Arg Ala 145 150 155

Ile Ala Gly Leu Ser Met Gly Gly Gly Gin Ser Phe Asn Ile Gly Leu 165 170 175

Thr Asn Leu Asp Lys Phe Ala Tyr Ile Gly Pro Ile Ser Ala Ala Pro 180 185 190

Asn Thr Tyr Pro Asn Glu Arg Leu Phe Pro Asp Gly Gly Lys Ala Ala 195 200 205

Arg Glu Lys Leu Lys Leu Phe Ile Ala Cys Gly Thr Asn Asp Ser

Leu Ile Gly Phe Gly Gln Arg Val His Glu Tyr Cys Val Ala Asn Asn 225 230 235 240

lle Asn His Val Tyr Trp Leu Ile Gln Gly Gly Gly His Asp Phe Asn 245 250 255

Val Trp Lys Pro Gly Leu Trp Asn Phe Leu Gln Met Ala Asp Glu Ala 260 265 270

Gly Leu Thr Arg Asp Gly Asn Thr Pro Val Pro Thr Pro Ser Pro Lys 275 280 285

Pro Ala Asn Thr Arg Ile Glu Ala Glu Asp Tyr Asp Gly Ile Asn Ser 290 295 300

Ser 305	Ser	Ile	Glu	Ile	Ile 310	Gly	Val	Pro	Pro	Glu 315	Gly	Gly	Arg	Gly	320	
Gly	Tyr	Ile	Thr	Ser 325	Gly	Asp	Туr	Leu	Val 330	Туr	Lys	Ser	Ile	Asp 335	Phe	
Gly	Asn	Gly	Ala 340	Thr	Ser	Phe	Lys	Ala 345	Lys	Val	Ala	Asn	Ala 350	Asn	Thr	
Ser	Asn	Ile 355	Glu	Leu	Arg	Leu	Asn 360	Gly	Pro	Asn	Gly	Thr 365	Leu	Ile	Gly	
Thr	Leu 370	Ser	Val	Lys	Ser	Thr 375	Gly	Asp	Trp	Asn	Thr 380	Tyr	Glu	Glu	Gln	
Thr 385	Cys	Ser	lle	Ser	Lys 390	Val	Thr	Gly	Ile	Asn 395	Asp	Leu	Tyr	Leu	Val 400	
Phe	Lys	G17	Pro	Vál 405	Asn	Ile	Asp	Trp	Phe 410	Thr	Phe	Gly	Val	Ser 415	Glu	
Lys	Asp	Glu	Leu 420													
<210 <210 <210 <210	1 > 2 >	109 1296 DNA arti		al s	eque	nce										
<220 <22		plas	mid	1126	7											
<400 atg	0> aggg	109 tgt	tgct	cgtt	gc c	ctcg	ctct	c ct	ggct	ctcg	ctg	cgag	cgc	cacc	agcgct	60
gcg	cagt	ccg	agcc	ggag	ct g	aagc	t gga	g tc	cgtg	gtga	tcg	tgtc	ccg	ccac	ggcgtg	120
cgc	gccc	cga	ccaa	ggcc	ac c	cagc	tcat	g ca	ggac	gtga	ccc	cgga	cgc	ctgg	ccgacc	180
tgg	cegg	tga	agct	caac	ga g	ctga	cccc	g cg	cggc	ggcg	agc	tgat	cgc	ctac	ctcggc	240
															ccgcag	300
															gaggcc	360
															acctcc	420
															gccaac	480
															cactac	540

cagaccgcct	tccgcgagct	ggagcgcgtg	ctcaacttcc	cgcagtccaa	cctctgcctc	600
aagcgcgaga	agcaggacga	gtcctgctcc	ctcacccagg	ccctcccgtc	cgagctgaag	660
gtgtccgccg	actgcgtgtc	cctcaccggc	gccgtgtccc	tcgcctccat	gctcaccgaa	720
atottectec	tccagcaggc	ccagggcatg	ccggagccgg	gctggggccg	catcaccgac	780
teccaccagt	ggaacaccct	cetetecete	cacaacgccc	agttcgacct	cctccagcgc	840
accccggagg	tggcccgctc	ccgcgccacc	ccgctcctcg	acctcatcaa	gaccgccctc	900
accccgcacc	coccocagaa	gcaggcctac	ggcgtgaccc	tcccgacctc	cgtgctcttc	960
atcgccggcc	acgacaccaa	cctcgccaac	ctcggcggcg	ccctggagct	gaactggacc	1020
ctcccgggcc	agccggacaa	caccccgccg	ggcggcga gc	tggtgttcga	gcgctggcgc	1080
egecteteeg	acaacteeca	gtggattcag	gtgtccctcg	tgttccagac	cctccagcag	1140
atgcgcgaca	agaccccgct	ctccctcaac	accccgccgg	gcgaggtgaa	gctcaccctc	1200
geeggetgeg	aggagcgcaa	cgcccagggc	atgtgctccc	tcgccggctt	cacccagatc	1260
		ggcctgctcc				1296

- <210> 110 <211> 431 <212> PRT <213> artificial sequence
- <223> plasmid 11267 aa sequence
- <400> 110

Met Arg Val Leu Leu Val Ala Leu Ala Leu Leu Ala Leu Ala Ser

Ala Thr Ser Ala Ala Gln Ser Glu Pro Glu Leu Lys Leu Glu Ser Val 20 25 30

Val Ile Val Ser Arg His Gly Val Arg Ala Pro Thr Lys Ala Thr Gln 40 35

Leu Met Gin Asp Val Thr Pro Asp Ala Trp Pro Thr Trp Pro Val Lys

Leu Gly Glu Leu Thr Pro Arg Gly Gly Glu Leu Ile Ala Tyr Leu Gly 70 75

His	Tyr	Trp	Arg	Gln	Arg	Leu	Val	Ala	Asp	Gly	Leu	Leu	Pro	Lys	Cys
	•	-	_	85					90					95	

- Gly Cys Pro Gln Ser Gly Gln Val Ala Ile Ile Ala Asp Val Asp Glu 100 105 110
- Arg Thr Arg Lys Thr Gly Glu Ala Phe Ala Ala Gly Leu Ala Pro Asp 115 120 125
- Cys Ala Ile Thr Val His Thr Gln Ala Asp Thr Ser Ser Pro Asp Pro 130 135 140
- Leu Phe Asn Pro Leu Lys Thr Gly Val Cys Gln Leu Asp Asn Ala Asn 145 150 155 160
- Val Thr Asp Ala Ile Leu Glu Arg Ala Gly Gly Ser Ile Ala Asp Phe 165 170 175
- Thr Gly His Tyr Gln Thr Ala Phe Arg Glu Leu Glu Arg Val Leu Asn 180 185 190
- Phe Pro Gln Ser Asn Leu Cys Leu Lys Arg Glu Lys Gln Asp Glu Ser 195 200 205
- Cys Ser Leu Thr Gln Ala Leu Pro Ser Glu Leu Lys Val Ser Ala Asp 210 215 220
- Cys Val Ser Leu Thr Gly Ala Val Ser Leu Ala Ser Met Leu Thr Glu 225 230 235 240
- Ile Phe Leu Leu Gln Gln Ala Gln Gly Met Pro Glu Pro Gly Trp Gly 245 250 255
- Arg Ile Thr Asp Ser His Gln Trp Asn Thr Leu Leu Ser Leu His Asn 260 270
- Ala Gln Phe Asp Leu Leu Gln Arg Thr Pro Glu Val Ala Arg Ser Arg 275 280 285
- Ala Thr Pro Leu Leu Asp Leu Ile Lys Thr Ala Leu Thr Pro His Pro 290 295 300
- Frc Gln Lys Gln Ala Tyr Gly Val Thr Leu Pro Thr Ser Val Leu Phe 305 310 315 320

lle Ala Gly His Asp Thr Ash Leu Ala Ash L 325 330	335
!eu Asn Trp Thr Leu Pro Gly Gln Pro Asp A 340 345	asn Thr Pro Pro Gly Gly 350
Glu Leu Val Phe Glu Arg Trp Arg Arg Leu S 355 360	er Asp Asn Ser Gln Trp 365
The Gln Val Ser Leu Val Phe Gln Thr Leu G 370 375	in Gln Met Arg Asp Lys 380
Thr Pro Leu Ser Leu Asn Thr Pro Pro Gly G 385 390 3	Slu Val Lys Leu Thr Leu 195 400
Ala Gly Cys Glu Glu Arg Asn Ala Gln Gly M 405 410	Met Cys Ser Leu Ala Gly 415
Phe Thr Gln Ile Val Asn Glu Ala Arg Ile P 420 425	Pro Ala Cys Ser Leu 430
<210> 111 <211> 1314 <212> DNA <213> artificial sequence	
<220> <223> plasmid 11268	
<400> 111 atgagggtgt tgctcgttgc cctcgctctc ctggctct	cg ctgcgagcgc caccagcgct 60
gcgcagtccg agccggaget gaagetggag teegtggt	ga togtgtocog coacggogtg 120
cgcgccccga ccaaggccac ccagctcatg caggacgt	ga ccccggacgc ctggccgacc 180
tggccggtga ageteggega getgaeceeg egeggegg	geg agetgatege etacetegge 240
cactactage gecagegeet egtggeegae ggeeteet	cc cgaagtgcgg ctgcccgcag 300
tccggccagg tggccatcat cgccgacgtg.gacgagcg	gca cccgcaagac cggcgaggcc 360
ttegeegeeg geetegeece ggaetgegee ateacegt	gc acacccaggc cgacacctcc 420
teeceggace egetetteaa eeegeteaag aceggegt	gt gecagetega caacgecaac 480
grgacegacg ccateergga gegegeegge ggetecat	cg ccgacttcac cggccactac 540
cagacogcot teegegaget ggagegegtg eteaactt	cc cgcagtccaa cctctgcctc 600

agcaggacga	gtcctgctcc	ctcacccagg	ccctcccgtc	cgagctgaag	660
actgcgtgtc	cctcaccggc	gccgtgtccc	tcgcctccat	gctcaccgaa	720
tccagcaggc	ccagggcatg	ccggagccgg	gctggggccg	catcaccgac	780
ggaacaccct	cctctccctc	cacaacgccc	agttcgacct	cctccagcgc	840
tggcccgctc	ccgcgccacc	ccgctcctcg	acctcatcaa	gaccgccctc	900
cgccgcagaa	gcaggcctac	ggcgtgaccc	tcccgacctc	cgtgctcttc	960
acgacaccaa	cctcgccaac	ctcggcggcg	ccctggagct	gaactggacc	1020
agccggacaa	caccccgccg	ggcggcgagc	tggtgttcga	gcgctggcgc	1080
acaactccca	gtggattcag	gtgtccctcg	tgttccagac	cctccagcag	1140
agaccccgct	ctccctcaac	accccgccgg	gcgaggtgaa	gctcaccctc	1200
aggagcgcaa	cgcccagggc	atgtgctccc	tcgccggctt	cacccagatc	1260
cccgcatccc	ggcctgctcc	ctctccgaga	aggacgagct	gtaa	1314
	actgcgtgtc tccagcaggc ggaacaccct tggcccgctc cgccgcagaa acgacaccaa agccggacaa acaactccca agaccccgct	actgcgtgtc cctcaccggc tccagcaggc ccagggcatg ggaacaccct cctctccetc tggccegctc ccgcgccacc cgccgcagaa gcaggcctac acgacaccaa cctcgccaac agccggacaa caccccgccg acaactccca gtggattcag agaccccgct ctccctcaac aggagcgcaa cgcccagggc	actgcgtgtc cctcaccggc gccgtgtccc tccagcaggc ccagggcatg ccggagccgg ggaacaccct cctccctc cacaacgccc tggcccgctc ccgcgccacc ccgctcctcg cgccgcagaa gcaggcctac ggcgtgaccc acgacaccaa cctcgccaac ctcggcggcg agccggacac gtggatcac gtggatcac acaactccca gtggattcag gtgtccctcg agacccgct ctccctcaac accccgccg aggagcgcaac acgacccgct ctccctcaac accccgccg aggagcgcaa cgcccaggc atgtgctccc	actgcgtgtc ceteaccgge geegtgteee tegeeteeat tecageagge ceagggcatg cegagecgg getggggeeg ggaacaccet ceteecte cacaacgee agttegacet tggeeggeeg cegegeegeegeegeegeegeegeegeegeegeegeegee	agcaggacga gtcctgctcc ctcaccagg ccctcccgtc cgagctgaag actgcgtgtc cctcaccgge gccgtgtccc tcgcctccat gctcaccgaa tccagcaggc ccagggcatg ccggagccgg gctggggccg catcaccgac ggaacaccct cctctcctc cacaacgccc agttcgacct cctccagcgc tggcccgctc ccgcgccacc ccgctcctcg acctcatcaa gaccgcctc acgaccacaa gccgcagaaa gcaggcctac ggcgtgaccc tcccgacctc cgtgctctc acgaccaca ctcggcagac tcccgaacct gaactggacc agccggacaa cacccagac ggcggcgagc tggtgttcga gcgctggcg acaactccca gtggattcag gtgtccctcg tgttccagac cctccagcag agaccccgct ctccctcaac acccgccgg gcgaggtgaa gctcaccctc aggagcgaa cgccagggc atgtgccc tcgccggct cacccagatc cacccagatc cccgcatccc ggcctgctcc ctctccgaga aggacgact gtaa

- <210> 112 <211> 437 <212> PRT <213> artificial sequence
- <223> plasmid 11268 amino acid sequence
- <400> 112

Met Arg Val Leu Leu Val Ala Leu Ala Leu Leu Ala Leu Ala Ala Ser

Ala Thr Ser Ala Ala Gln Ser Glu Pro Glu Leu Lys Leu Glu Ser Val

Val Ile Val Ser Arg His Gly Val Arg Ala Pro Thr Lys Ala Thr Gln 35

Leu Met Gln Asp Val Thr Pro Asp Ala Trp Pro Thr Trp Pro Val Lys

Leu Gly Glu Leu Thr Pro Arg Gly Gly Glu Leu Ile Ala Tyr Leu Gly

His Tyr Trp Arg Gin Arg Leu Val Ala Asp Gly Leu Leu Pro Lys Cys

Page 133

- Gly Cys Pro Gln Ser Gly Gln Val Ala Ile Ile Ala Asp Val Asp Glu 100 105 110
- Arg Thr Arg Lys Thr Gly Glu Ala Phe Ala Ala Gly Leu Ala Pro Asp 115 125
- Cys Ala Ile Thr Val His Thr Gln Ala Asp Thr Ser Ser Pro Asp Pro 130 135 140
- Leu Phe Asn Pro Leu Lys Thr Gly Val Cys Gln Leu Asp Asn Ala Asn 145 150 155 160
- Val Thr Asp Ala Ile Leu Glu Arg Ala Gly Gly Ser Ile Ala Asp Phe 175
- Thr Gly His Tyr Gln Thr Ala Phe Arg Glu Leu Glu Arg Val Leu Asn 180 185 190
- Phe Pro Gln Ser Asn Leu Cys Leu Lys Arg Glu Lys Gln Asp Glu Ser 195 200 205
- Cys Ser Leu Thr Gln Ala Leu Pro Ser Glu Leu Lys Val Ser Ala Asp 210 215 220
- Cys Val Ser Leu Thr Gly Ala Val Ser Leu Ala Ser Met Leu Thr Glu 225 230 236
- Ile Phe Leu Gin Gin Ala Gin Gly Met Pro Glu Pro Gly Trp Gly 245 250 255
- Arg lle Thr Asp Ser His Gln Trp Asn Thr Leu Leu Ser Leu His Asn 260 265 270
- Ala Gln Phe Asp Leu Leu Gln Arg Thr Pro Glu Val Ala Arg Ser Arg 275 280 285
- Ala Thr Pro Leu Leu Asp Leu Ile Lys Thr Ala Leu Thr Pro His Pro 290 295 300
- Pro Gln Lys Gln Ala Tyr Gly Val Thr Leu Pro Thr Ser Val Leu Phe 305 310 315

- Ile Ala Gly His Asp Thr Asn Leu Ala Asn Leu Gly Gly Ala Leu Glu 325 330 335
- Glu Leu Val Phe Glu Arg Trp Arg Arg Leu Ser Asp Asn Ser Gln Trp 355 360 365
- Ile Gln Val Ser Leu Val Phe Gln Thr Leu Gln Gln Met Arg Asp Lys 370 380
- Thr Pro Leu Ser Leu Asn Thr Pro Pro Gly Glu Val Lys Leu Thr Leu 385 390 395
- Ala Gly Cys Glu Glu Arg Asn Ala Gln Gly Met Cys Ser Leu Ala Gly 405 410 415
- Phe Thr Gln Ile Val Asn Glu Ala Arg Ile Pro Ala Cys Ser Leu Ser 420 425 430
- Glu Lys Asp Glu Leu 435